

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:00:53 ; Search time 20.9296 Seconds
(without alignments)
757.625 Million cell updates/sec

Title: US-09-627-165E-14

Perfect score: 596

Sequence: 1 ARNPPIXMLRQINGCESS.....VVDANDVTCTXSEPTVRIV 119

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	590	99.0	120	AA847096	
2	413.5	69.4	551	AB879450	Korean mistletoe 1
3	408.5	68.5	174	AB847097	Galactose-recogins
4	387	64.9	564	AA10021	Korean mistletoe 1
5	387	64.9	564	AAW90127	Prepro mistletoe 1
6	386	64.8	531	AA25979	Mistletoe lectin p
7	386	64.8	532	AA25982	Mistletoe lectin i
8	373	62.6	533	AA25970	Mistletoe lectin p
9	373	62.6	533	AA25973	Mistletoe lectin p
10	373	62.6	533	AA25976	Mistletoe lectin p

11	327	54.9	256	20	AA25981	Mistletoe lectin A
12	327	54.9	256	20	AA25984	Mistletoe lectin A
13	317	53.2	254	20	AA25980	Mistletoe lectin A
14	317	53.2	254	20	AA25983	Mistletoe lectin A
15	311	52.2	252	19	AAW64659	Mistletoe rMLA pro
16	311	52.2	252	19	AAW64661	Mistletoe rMLA var
17	311	52.2	253	18	AAW10022	Prepro mistletoe 1
18	311	52.2	253	20	AAW90125	Mistletoe MU A-cha
19	304	51.0	253	20	AA25971	Mistletoe lectin A
20	304	51.0	255	20	AA25974	Mistletoe lectin A
21	304	51.0	255	20	AA25977	Mistletoe lectin A
22	303	50.8	254	22	AA847090	A-chain isoform fo
23	298	50.0	254	22	AA847091	A-chain isoform fo
24	298	50.0	256	22	AA847092	A-chain isoform fo
25	178	29.9	332	8	AA70097	Ricin A. Escheric
26	178	29.9	332	8	AA70838	Sequence of Ricinu
27	178	29.9	332	10	AA95639	Ricin A encoded by
28	178	29.9	332	11	AA80654	Ricin A gene produ
29	178	29.9	562	10	AA90079	Ricin D. Ricinus
30	178	29.9	565	6	AA50166	Sequence of prepro
31	178	29.9	565	7	AA60240	PreproRicinu. Ricl
32	178	29.9	565	22	AA678300	Castor bean prepro
33	178	29.9	574	8	AA70325	Sequence of Ricinu
34	178	29.9	574	10	AA94793	DNA sequence of rl
35	178	29.9	576	8	AA70326	Sequence of Ricinu
36	178	29.9	576	18	AAW25787	Castorbean ricin.
37	178	29.9	576	20	AAV55892	Castor bean ricin
38	178	29.9	576	21	AAV78592	Ricinus communis r
39	178	29.9	576	22	AA678301	Castor bean prepro
40	178	29.9	576	22	AA678302	Castor bean prepro
41	177	29.7	280	10	AA95648	Ricin agglutinin A
42	177	29.7	540	18	AAW25143	Castor oil plant a
43	177	29.7	540	18	AAW21706	R. communis agglut
44	174	29.2	534	8	AA70324	Sequence of Ricin
45	165	21.7	565	22	AA678304	Modified castor be

ALIGNMENTS

RESULT 1	
AA847096	
ID	AA847096 standard; Protein: 120 AA.
XX	
AC	AA847096;
XX	
DT	16-MAY-2001 (first entry)
XX	
DE	Korean mistletoe lectin #1.
XX	
KW	Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KW	KML; tumour; KM-110; KML-C; KMHP; KML-IIU; KML-IL;
XX	heparin binding protein.
XX	
OS	Viscum album coloratum.
XX	
FH	Key
FT	Misc-difference 7 Location/Qualifiers
FT	Misc-difference 19 /note= "Encoded by NTG"
FT	Misc-difference 19 /note= "Encoded by TCN"
FT	Misc-difference 64 /note= "Encoded by NTG"
FT	Misc-difference 111 /note= "Encoded by NTG"
FT	Misc-difference 111 /note= "Encoded by NTG"
PN	EP1074560-A2.
XX	
PD	07-FEB-2001.
XX	
PF	27-JUL-2000; 2000EP-0402163.
XX	
PR	27-JUL-1999; 99KR-0030638.

```

XX (MIST-) MISTLE BIOTECH CO LTD.
PA
XX
PI Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI Kang T, Park C;
XX
DR WPI: 2001-171044/18.
DR N-PSDB; AAC47096.
XX
PT Novel lectin proteins isolated from Korean mistletoe, useful for
PT enhancing immunity and effectuating anti-tumoral activity -
PS Claim 31; Page 33; 62pp; English.
XX
CC The sequences given in AAB47096-97 are lectins isolated from Korean
CC mistletoe. Korean mistletoe lectins (KML) are useful for enhancing
CC immunity and for treating tumours. The KML's are isolated from a
CC protein fraction derived from the leaves, stems and fruits of Korean
CC mistletoe, which is designated KM-110. One of the isolates, KML-C was
CC shown to be effective against colon 26-M3.1 carcinoma and L5178Y-M25
CC lymphoma.
CC
XX Sequence 120 AA;
SO
Query Match 99.0%; Score 590; DB 22; Length 120;
Best Local Similarity 100.0%; Pred. No. 8,7e-66;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ARFNPXMRRLKROINGESSPPNMYMLELTSWGRSTQVQSKDGIENQIRLQISAGN 60
DB 1 ARFNPXMRRLKROINGESSPPNMYMLELTSWGRSTQVQSKDGIENQIRLQISAGN 60
QY 61 FVTXSNVRDVISSLAIMLFECGSRPFSSLDHPSPLLRSVVDANDVCTXSEPTVIR 119
DB 61 FVTXSNVRDVISSLAIMLFECGSRPFSSLDHPSPLLRSVVDANDVCTXSEPTVIR 119
RESULT 2
AAB79450
ID AAB79450 standard; Protein; 551 AA.
XX
AC AAB79450;
XX
DT 08-JUL-2002 (first entry)
XX
DE Galactose-recognising mistletoe lectin.
XX
KM Mistletoe; galactose-recognising mistletoe lectin; MLIII.
XX
OS Viscum album.
XX
FH Key Location/Qualifiers
FT Misc-difference 223 /note= "Encoded by ATG"
FT Misc-difference 251 /note= "Encoded by TTT"
FT Misc-difference 344 /note= "Encoded by TCG"
FT Misc-difference 380 /note= "Encoded by GCC"
FT Misc-difference 448 /note= "Encoded by GTG"
FT Misc-difference 448 /note= "Encoded by GTG"
XX
PN DE10044027-A1.
XX
PD 14-MAR-2002.
XX
PF 06-SEP-2000; 2000DE-1044027.
XX
PR 06-SEP-2000; 2000DE-1044027.
XX
PA (VISC-) VISCUM AG.
XX

```

```

PI Kleff S;
XX
DR WPI: 2002-316737/36.
DR N-PSDB; ABL56947.
XX
PT New nucleic acid encoding preprotein of mistletoe lectin, useful as
PT diagnostic and therapeutic agents, also encodes polypeptide -
PS Claim 1; Fig 1; 6pp; German.
XX
CC The invention relates to a nucleic acid molecule (ABL56947) that encodes
CC a preprotein (AAB79450) which, after maturation, has the biological
CC activity of the galactose-recognising mistletoe lectin (MLIII). The MLIII
CC encoding nucleic acid molecule, primers specific to it or complements of
CC it, and encoded (oligomeric) polypeptides are useful as diagnostic and
CC therapeutic agents.
XX
SO Sequence 551 AA;
SO
Query Match 69.4%; Score 413.5; DB 23; Length 551;
Best Local Similarity 69.7%; Pred. No. 8,3e-43;
Matches 85; Conservative 12; Mismatches 22; Indels 3; Gaps 1;
QY 1 ARFNPXMRRLKROINGESSPPNMYMLELTSWGRSTQVQSKDGIENQIRLQISAGN 60
DB 200 ARFNPXMRRLKROINGESSPPNMYMLELTSWGRSTQVQSKDGIENQIRLQISAGN 259
QY 61 FVTXSNVRDVISSLAIMLFECGSRPFSSLDHPSPLLRSVVD--ANDVCTXSEPTVIR 117
DB 260 FVTXSNVRDVISSLAIMLFECGSRPFSSLDHPSPLLRSVVD--ANDVCTXSEPTVIR 319
QY 118 IV 119
DB 320 IV 321
RESULT 3
AAB47097
ID AAB47097 standard; Protein; 174 AA.
XX
AC AAB47097;
XX
DT 16-MAR-2001 (first entry)
XX
DE Korean mistletoe lectin #2.
XX
KM Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KM KML; tumour; KM-110; KML-C; KMBP; KML-III;
KM heparin binding protein.
XX
OS Viscum album coloratum.
XX
FH EP1074560-A2.
FT 07-FEB-2001.
PF 27-JUL-2000; 2000EP-0402168.
PR 27-JUL-1999; 99KR-0030638.
XX
PA (MIST-) MISTLE BIOTECH CO LTD.
XX
PI Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI Kang T, Park C;
XX
DR WPI: 2001-171044/18.
DR N-PSDB; AAC85479.
XX
PT Novel lectin proteins isolated from Korean mistletoe, useful for
PT enhancing immunity and effectuating anti-tumoral activity -
XX
PS Claim 32; Page 34; 62pp; English.
XX

```


[illegible]

DE	Mistletoe lectin I (variant) protein fragment.
KX	Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW	lymphokine 28S subunit; non-cytotoxic; T-cell activation; immune response;
KM	lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW	cancer; cytotoxicity; antigen; isoform; lectin I.
XX	
OS	Viscum album.
XX	
PN	DEL9804210-A1.
XX	
PD	12-AUG-1999.
XX	
PF	03-FEB-1998; 98DE-1004210.
XX	
XX	03-FEB-1998; 98DE-1004210.
XX	
PA	(BIOS-) BIOSYN ARZNEIMITTEL GMBH.
PI	Morris P, Stiefel T, Voelter W, Welters P,
XX	
DR	WPL; 1999-445335/38.
XX	
N-PDSB; AAZ09106.	
PT	Preparation of mistletoe lectins in heterologous systems,
PS	particularly for use as anticancer agents and immunostimulants
XX	
XX	Disclosure; Fig 4B; 78pp; German.
CC	This invention describes a novel mistletoe lectin (I) and its fragments
CC	which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC	of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC	ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC	lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC	fragments are used to treat uncontrolled cell growth (particularly
CC	cancers) and if they lack cytotoxicity, to increase the strength of the
CC	immune response, particularly to a co-administered antigen
CC	(tumour-associated, bacterial or viral). The method allows production of
CC	mistletoe lectin, and its individual chains, in many different isoforms
CC	and on a large scale, at any time of the year. Recombinant products are
CC	free from toxins present in natural mistletoe extracts. This sequence
CC	represents a fragment of a mistletoe lectin I protein variant.
XX	
SQ	Sequence 532 AA:
Query Match	64.8%; Score 386; DB 20; Length 532;
Best Local Similarity	67.2%; Pred. No. 2.2e-39;
Matches 80; Conservative 12; Mismatches 25; Indels 2; Gaps 1.	
QY	1 ARENPILKRLRROINSSESPNNMYALELETSMGSGOSTGYVOQSXDGIINTQIRLQISAGN 60 : : : :
Db	167 ARNPILMRARROYINGASFLPDDVYMLEELTSSWGQSTGYOVSHTDYGFNPRILALPPGN 226
QY	61 FVTXSNVRDYISSLAIMEFECSGRPSSLDHPSPLLLRSVDAANDVCTGXSEPTVRIY 119 : : : : : : : : : :
Db	227 FVTLTNVRDYIASLAIMLFVCGERPSSSDVRYPVLIVPI--ADVDTCASASEPTVRIY 283
RESULT 8	
AAV25970	standard; protein; 533 AA.
ID	AAV25970;
AC	AAV25970;
DT	18-OCT-1999 (first entry)
XX	
DE	Mistletoe lectin protein consensus sequence 1.
XX	
KW	Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KW	lymphokine 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW	lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW	cancer; cytotoxicity; antigen; isoform.
XX	


```

OS Viscum album.
XX
FH Key Location/Qualifiers
FT Misc-difference 15
FT /label= Asp, Glu
FT Misc-difference 63
FT /label= Gly, Gln
FT Misc-difference 66
FT /label= Ile, Val
FT Misc-difference 75
FT /label= Leu, Ala
FT Misc-difference 107
FT /label= Asp, Arg or none
FT Misc-difference 113
FT /label= Asn, Thr
FT Misc-difference 117
FT /label= Pro, Thr
FT Misc-difference 134
FT /label= Asp, Glu
FT Misc-difference 141
FT /label= Ser, Thr
FT Misc-difference 145
FT /label= Phe, Tyr
FT Misc-difference 152
FT /label= Thr, Ala
FT Misc-difference 177
FT /label= Ala, Tyr
FT Misc-difference 180
FT /label= Tyr, Asp
FT Misc-difference 185
FT /label= Ala, Glu
FT Misc-difference 191
FT /label= Val, Met
FT Misc-difference 219
FT /label= Ile, Phe
FT Misc-difference 224
FT /label= Pro, Ser
FT Misc-difference 225
FT /label= Pro, Thr
FT Misc-difference 232
FT /label= Thr, Ser
FT Misc-difference 236
FT /label= Asp, Ser
FT Misc-difference 287
FT /label= Asn, Ser
FT Misc-difference 290
FT /label= Cys, Arg
FT Misc-difference 325
FT /label= Gly, Asn
FT Misc-difference 364
FT /label= Gly, Asp
FT Misc-difference 426
FT /label= Gly, Gln
FT Misc-difference 435
FT /label= Val, Asp
FT Misc-difference 439
FT /label= Gln, Lys
FT Misc-difference 442
FT /label= Gly or none
FT Misc-difference 443
FT /label= Arg, Lys
FT Misc-difference 464
FT /label= Cys, Ser, Val
FT Misc-difference 480
FT /label= Ala, Gly
FT Misc-difference 481
FT /label= Gly, Ala
FT Misc-difference 483
FT /label= Ser, Gly
FT Misc-difference 484
FT /label= Gly, Ser
FT Misc-difference 493
FT /label= Gly, Tyr

```

```

FT Misc-difference 500
FT /label= Asn, Ser, Thr, Lys
FT Misc-difference 501
FT /label= Ser, Gly
FT Misc-difference 502
FT /label= Leu, Pro
FT Misc-difference 503
FT /label= Ala, Met
FT Misc-difference 504
FT /label= Met, Val
FT Misc-difference 533
FT /label= Pro, Phe
XX
PN DE19804210-A1.
XX
XX 12-AUG-1999.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX Morris P, Stiefel T, Voelter W, Walters P;
XX MPI; 1999-445335/38.
XX
XX Claim 1; Page 25-26; 78pp; German.
XX
CC This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains. In many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a consensus sequence of the mistletoe lectin described in the
CC specification.
XX
SQ Sequence 533 AA:
XX
XX Query Match 62.6%; Score 373; DB 20; Length 533;
XX Best Local Similarity 65.5%; Pred. No. 9.3e-38;
XX Matches 78; Conservative 10; Mismatches 29; Indels 2; Gaps 1;
XX
OY 1 ARENPIMRLRQINSSESPNNMYMLEETSMGROSYOVQOSKDFNTQIRLOISGN 60
DB 168 ARENPIMLRXQKINSXGFLPDXYMLEETSMGQOSYOVQSTGCVNRPRLIXXGN 227
OY 61 FVTXSNVDYVYSSLAIMLEFGSGRPSSLDHPSPILLNSVDANDVCTCTSEPTVRIV 119
DB 228 FVTLKNVRXYIASLAIMLEFVCGERSSSDVRYRPIVRVI--ADDVYCSASEPTVRIV 284
XX
XX RESULT 9
XX AAY25973
XX ID AAY25973 standard; protein: 533 AA.
XX
XX AAY25973;
XX
XX 18-OCT-1999 (first entry)
XX
XX Mistletoe lectin protein consensus sequence 2.
XX
XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;

```

KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 XX cancer; cytotoxicity; antigen; isoform.
 OS Viscum album.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 15
 FT /label= Asp, Glu
 FT Misc-difference 63
 FT /label= Gly, Gln
 FT Misc-difference 66
 FT /label= Ile, Val
 FT Misc-difference 75
 FT /label= Leu, Ala
 FT Misc-difference 107
 FT /label= Asp, Arg, none
 FT Misc-difference 113
 FT /label= Asn, Thr
 FT Misc-difference 117
 FT /label= Pro, Thr
 FT Misc-difference 134
 FT /label= Asp, Glu
 FT Misc-difference 141
 FT /label= Ser, Thr
 FT Misc-difference 145
 FT /label= Phe, Tyr
 FT Misc-difference 152
 FT /label= Thr, Ala
 FT Misc-difference 177
 FT /label= Ala, Tyr
 FT Misc-difference 180
 FT /label= Tyr, Asp
 FT Misc-difference 185
 FT /label= Ala, Glu
 FT Misc-difference 191
 FT /label= Val, Met
 FT Misc-difference 219
 FT /label= Ile, Phe
 FT Misc-difference 224
 FT /label= Phe, Ser
 FT Misc-difference 225
 FT /label= Pro, Thr
 FT Misc-difference 232
 FT /label= Thr, Ser
 FT Misc-difference 236
 FT /label= Asp, Ser
 FT Misc-difference 287
 FT /label= Asn, Ser
 FT Misc-difference 290
 FT /label= Cys, Arg
 FT Misc-difference 325
 FT /label= Gly, Asn
 FT Misc-difference 364
 FT /label= Gly, Asp
 FT Misc-difference 427
 FT /label= Gly, Gln
 FT Misc-difference 435
 FT /label= Val, Asp
 FT Misc-difference 439
 FT /label= Gln, Lys
 FT Misc-difference 442
 FT /label= Gly, none
 FT Misc-difference 443
 FT /label= Arg, Lys
 FT Misc-difference 464
 FT /label= Cys, Ser, Val
 FT Misc-difference 480
 FT /label= Ala, Gly
 FT Misc-difference 481
 FT /label= Gly, Ala
 FT Misc-difference 483
 FT /label= Ser, Gly

FT Misc-difference 484
 FT /label= Gly, Ser
 FT Misc-difference 493
 FT /label= Gly, Tyr
 FT Misc-difference 500
 FT /label= Asn, Ser, Thr, Lys
 FT Misc-difference 501
 FT /label= Ser, Gly
 FT Misc-difference 502
 FT /label= Leu, Pro
 FT Misc-difference 503
 FT /label= Ala, Met
 FT Misc-difference 504
 FT /label= Met, Val
 FT Misc-difference 533
 FT /label= Pro, Phe
 XX
 PN DE19804210-A1.
 XX
 XX 12-AUG-1999.
 PD
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI; 1999-445335/38.
 DR
 XX
 XX Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 PS Claim 4; Page 28-29; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of the mistletoe lectin described in the
 CC specification.
 CC
 SQ Sequence 533 AA;
 XX
 Query Match 62.6%; Score 373; DB 20; Length 533;
 Best Local Similarity 65.5%; Pred. No. 9, 3e-38;
 Matches 78; Conservative 10; Mismatches 29; Indels 2; Gaps 1;
 QY 1 ARENPXWRLRQINSSESPNNMYLETSGROSTOVQOSQIDFNTQIRLQISAGN 60
 DB 168 ARFNPIIMRXRXKXINSXSFLLPDXYMLETSGOOSTGVQSHSTDFVFNXPRLAIXXGN 227
 QY 61 FVYXSNRVDYISSLAIMLFECSGRPSSLDHPSPLLRSVYDANVYCTXSPPYRIY 119
 DB 228 FVTLXNVXVIASIALIMLFVCGERPSSDVRYWPLVIRPI--ADDVTCASBPYRIY 284
 RESULT 10
 AAY25976
 ID AAY25976 standard; protein; 533 AA.
 XX
 AC AAY25976;
 XX
 DT 18-OCT-1999 (first entry)

```

XX Mistletoe lectin protein consensus sequence 3.
DE Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
XX Mistletoe; 28S subunit; non-cytotoxic; T-cell activation; immune response;
KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KM cancer; cytotoxicity; antigen; isoform.
XX
OS Viscum album.
XX
FH Key Location/Qualifiers
FT Misc-difference 15 /label= Asp, Glu
FT Misc-difference 63 /label= Gly, Gln
FT Misc-difference 66 /label= Ile, Val
FT Misc-difference 75 /label= Leu, Ala
FT Misc-difference 107 /label= Asp, none
FT Misc-difference 113 /label= Asn, Thr
FT Misc-difference 117 /label= Pro, Thr
FT Misc-difference 134 /label= Asp, Glu
FT Misc-difference 141 /label= Ser, Thr
FT Misc-difference 145 /label= Phe, Tyr
FT Misc-difference 152 /label= Thr, Ala
FT Misc-difference 177 /label= Ala, Tyr
FT Misc-difference 180 /label= Tyr, Asp
FT Misc-difference 185 /label= Ala, Glu
FT Misc-difference 191 /label= Val, Met
FT Misc-difference 219 /label= Ile, Phe
FT Misc-difference 224 /label= Pro, Ser
FT Misc-difference 225 /label= Pro, Thr
FT Misc-difference 232 /label= Thr, Ser
FT Misc-difference 236 /label= Asp, Ser
FT Misc-difference 287 /label= Asn, Ser
FT Misc-difference 290 /label= Cys, Arg
FT Misc-difference 325 /label= Gly, Asn
FT Misc-difference 364 /label= Gly, Asp
FT Misc-difference 426 /label= Gly, Gln
FT Misc-difference 435 /label= Val, Asp
FT Misc-difference 439 /label= Gln, Lys
FT Misc-difference 442 /label= Gly, none
FT Misc-difference 443 /label= Arg, Lys
FT Misc-difference 464 /label= Cys, Ser, Val
FT Misc-difference 480 /label= Ala, Gly
FT

```

```

FT Misc-difference 481 /label= Gly, Ala
FT Misc-difference 483 /label= Ser, Gly
FT Misc-difference 484 /label= Gly, Ser
FT Misc-difference 493 /label= Gly, Tyr
FT Misc-difference 500 /label= Asn, Ser, Thr, Lys
FT Misc-difference 501 /label= Ser, Gly
FT Misc-difference 502 /label= Leu, Pro
FT Misc-difference 503 /label= Ala, Met
FT Misc-difference 504 /label= Met, Val
FT Misc-difference 533 /label= Pro, Phe
FT
FT
PN DE19804210-A1.
XX
PD 12-AUG-1999.
XX
PF 03-FEB-1998; 98DE-1004210.
XX
PR 03-FEB-1998; 98DE-1004210.
XX
PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
PI Morris P, Stiefel T, Voelter W, Welters P;
XX WPI; 1999-445335/38.
XX
PT Preparation of mistletoe lactins in heterologous systems,
PS particularly for use as anticancer agents and immunostimulants
XX
PS Claim 40; Page 37-38; 78pp; German.
XX
CC This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a consensus sequence of the mistletoe lectin described in the
CC specification.
XX
SQ Sequence 533 AA:
SQ
Query Match 62.6%; Score 373; DB 20; Length 533;
Best Local Similarity 65.5%; Pred No. 9 3e-38;
Matches 78; Conservative 10; Mismatches 29; Indels 2; Gaps 1;
QY 1 ARENPXWRLRQINSSESPNMYMLETSGROSTOVQOSKIDIFMTQIRLQISAGN 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 168 ARFNPILMRXQXNSGXSLPDXMYMLETSGQOSTOVGHSTGVFNNPKRLAIXGN 227
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 FVTXSNRVYISSLAIMLFCSCGRPFSSLDHPSPLLRSVDANANVTCTSPYRIY 119
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 228 FVTLXNVRYIASLAIMLFCVGERPSSSDVRYWPLVIRPI--ADDVTCASAEPTRIY 284
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 11
AAI25981
ID AAI25981 standard; Protein; 256 AA.

```

```

XX AC AAY25981;
XX XX 18-OCT-1999 (first entry)
XX DT
XX DE Mistletoe lectin A2 protein fragment.
XX KM Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
XX KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX KW cancer; cytotoxicity; antigen; isoform; lectin A2.
XX OS
XX PN Viscum album.
XX PD DEL9804210-A1.
XX XX 12-AUG-1999.
XX PF 03-FEB-1998; 98DE-1004210.
XX PR 03-FEB-1998; 98DE-1004210.
XX PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX PI Morris P, Stiefel T, Voelter W, Welters P;
XX DR WPI; 1999-445335/38.
XX DR N-PSDB; AAZ09105.
XX PT Preparation of mistletoe lectins in heterologous systems,
XX PT particularly for use as anticancer agents and immunostimulants
XX PS Claim 8; Fig 3B; 78pp; German.
XX XX
XX CC This invention describes a novel mistletoe lectin (I) and its fragments
XX CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
XX CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX CC fragments are used to treat uncontrolled cell growth (particularly
XX CC cancers) and if they lack cytotoxicity, to increase the strength of the
XX CC immune response, particularly to a co-administered antigen
XX CC (tumour-associated bacterial or viral). The method allows production of
XX CC mistletoe lectin, and its individual chains, in many different isoforms
XX CC and on a large scale, at any time of the year. Recombinant products are
XX CC free from toxins present in natural mistletoe extracts. This sequence
XX CC represents a fragment of the mistletoe lectin A2 protein.
XX SQ
XX
XX Query Match 54.9%; Score 327; DB 20; Length 256;
XX Best Local Similarity 73.9%; Pred. No. 2e-32;
XX Matches 65; Conservative 5; Mismatches 18; Indels 0; Gaps 0;
XX
XX QY 1 ARENPIMRLRQINSSESPNNMYMLETSMGRSTOVQOSKDGIFNTQIRLQISAGN 60
XX DB 169 ARENPIMRLRQINSSESPNNMYMLETSMGRSTOVQOSKDGIFNTQIRLQISAGN 228
XX QY 61 FVTXSNVRDVISSLAIMLFECSGRPFSS 88
XX DB 229 FVTLSNVRSVIASLAIMLFECSGRPFSS 256
XX
XX RESULT 12
XX AAY25984
XX ID AAY25984 standard; Protein; 256 AA.
XX AC AAY25984;
XX XX
XX DT 18-OCT-1999 (first entry)
XX DE Mistletoe lectin A2 (variant) protein fragment.
XX XX

```

```

XX KM Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
XX KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX KW cancer; cytotoxicity; antigen; isoform; lectin A2.
XX OS
XX PN Viscum album.
XX PD DEL9804210-A1.
XX XX 12-AUG-1999.
XX PF 03-FEB-1998; 98DE-1004210.
XX PR 03-FEB-1998; 98DE-1004210.
XX PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX PI Morris P, Stiefel T, Voelter W, Welters P;
XX DR WPI; 1999-445335/38.
XX DR N-PSDB; AAZ09108.
XX PT Preparation of mistletoe lectins in heterologous systems,
XX PT particularly for use as anticancer agents and immunostimulants
XX PS Disclosure; Fig 6B; 78pp; German.
XX XX
XX CC This invention describes a novel mistletoe lectin (I) and its fragments
XX CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
XX CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX CC fragments are used to treat uncontrolled cell growth (particularly
XX CC cancers) and if they lack cytotoxicity, to increase the strength of the
XX CC immune response, particularly to a co-administered antigen
XX CC (tumour-associated bacterial or viral). The method allows production of
XX CC mistletoe lectin, and its individual chains, in many different isoforms
XX CC and on a large scale, at any time of the year. Recombinant products are
XX CC free from toxins present in natural mistletoe extracts. This sequence
XX CC represents a fragment of a mistletoe lectin A2 protein variant.
XX SQ
XX
XX Query Match 54.9%; Score 327; DB 20; Length 256;
XX Best Local Similarity 73.9%; Pred. No. 2e-32;
XX Matches 65; Conservative 5; Mismatches 18; Indels 0; Gaps 0;
XX
XX QY 1 ARENPIMRLRQINSSESPNNMYMLETSMGRSTOVQOSKDGIFNTQIRLQISAGN 60
XX DB 169 ARENPIMRLRQINSSESPNNMYMLETSMGRSTOVQOSKDGIFNTQIRLQISAGN 228
XX QY 61 FVTXSNVRDVISSLAIMLFECSGRPFSS 88
XX DB 229 FVTLSNVRSVIASLAIMLFECSGRPFSS 256
XX
XX RESULT 13
XX AAY25980
XX ID AAY25980 standard; Protein; 254 AA.
XX AC AAY25980;
XX XX
XX DT 18-OCT-1999 (first entry)
XX DE Mistletoe lectin A1 protein fragment.
XX XX
XX KM Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
XX KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX KW cancer; cytotoxicity; antigen; isoform; lectin A1.
XX OS
XX Viscum album.
XX XX

```

PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX
 DR WPI: 1999-445335/38.
 XX
 DR N-PSDB; AAZ09104.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 XX
 PS Disclosure; Fig 2B; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of the mistletoe lectin A1 protein.
 XX
 SQ Sequence 254 AA:
 XX
 Query Match 53.2%; Score 317; DB 20; Length 254;
 Best Local Similarity 71.6%; Pred. No. 3.5e-31;
 Matches 63; Conservative 7; Mismatches 18; Indels 0; Gaps 0;
 QY 1 ARFNPIXRRLRQINSSESPNNMYMLETSMGROSTQVOQSKGIFNTQIRLOISAGN 60
 Db 167 ARFNPIXRRLRQINSSESPNNMYMLETSMGROSTQVOHSTQVFNPNIRLAIPGN 226
 QY 61 FVTXSNVRDVISSLAIMLFECGSRPSS 88
 Db 227 FVTLTNVRDIASLAIMLFVCGSRPSS 254
 XX
 RESULT 14
 AAY25983
 ID AAY25983 standard; Protein: 254 AA.
 XX
 AC AAY25983;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Mistletoe lectin A1 (variant) protein fragment.
 XX
 OS Mistletoe lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribosome; 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A1.
 XX
 OS Viscum album.
 XX
 PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.

XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX
 DR WPI: 1999-445335/38.
 XX
 DR N-PSDB; AAZ09107.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 XX
 PS Disclosure; Fig 5B; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of a mistletoe lectin A1 protein variant.
 XX
 SQ Sequence 254 AA:
 XX
 Query Match 53.2%; Score 317; DB 20; Length 254;
 Best Local Similarity 71.6%; Pred. No. 3.5e-31;
 Matches 63; Conservative 7; Mismatches 18; Indels 0; Gaps 0;
 QY 1 ARFNPIXRRLRQINSSESPNNMYMLETSMGROSTQVOQSKGIFNTQIRLOISAGN 60
 Db 167 ARFNPIXRRLRQINSSESPNNMYMLETSMGROSTQVOHSTQVFNPNIRLAIPGN 226
 QY 61 FVTXSNVRDVISSLAIMLFECGSRPSS 88
 Db 227 FVTLTNVRDIASLAIMLFVCGSRPSS 254
 XX
 RESULT 15
 AAW64659
 ID AAW64659 standard; Protein: 252 AA.
 XX
 AC AAW64659;
 XX
 DT 23-OCT-1998 (first entry)
 XX
 DE Mistletoe rMLA protein.
 XX
 OS Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; ricin; translocation.
 XX
 OS Viscum album.
 XX
 PN WO9829540-A2.
 XX
 PD 09-JUL-1998.
 XX
 PF 02-JAN-1998; 98WO-EP00009.
 XX
 PR 02-JAN-1997; 97EP-0100012.
 XX
 PA (BRAT-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
 XX
 PI Eck J, Schmidt A, Zinke H;
 XX
 DR WPI: 1998-388122/33.
 XX
 DR N-PSDB; AAV51341.


```

; CURRENT FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/022273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; EARLIER FILING DATE: 1995-06-26
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Viscum album
us-08-776-059-39

Query Match          52.2%; Score 311; DB 4; Length 235;
Best Local Similarity 71.8%; Pred. No. 1.9e-32;
Matches 61; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 1 ARFNPIXRRLROINSSESPNNMTELETSMGRQSTOVQOOSKDGFINTQIRIQISAGN 60
DB 150 ARFNPIXRRLROINSSESPNNMTELETSMGRQSTOVQOOSKDGFINTQIRIQISAGN 209
QY 61 FVTXSNVRDVYSSLAIMLFECGSRP 85
DB 210 FVTXSNVRDVYSSLAIMLFECGSRP 234

RESULT 3
us-08-776-059-31
; Sequence 31, Application US/08776059B
; Patent No. 6271368
; GENERAL INFORMATION:
; APPLICANT: LENTZEN, Hans
; APPLICANT: ECK, Jurgen
; APPLICANT: BAUR, Axel
; APPLICANT: ZINKE, Holger
; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
; FILE REFERENCE: 674503-2003
; CURRENT APPLICATION NUMBER: US/08/776,059B
; CURRENT FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/02273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; EARLIER FILING DATE: 1995-06-26
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Viscum album
us-08-776-059-31

Query Match          52.2%; Score 311; DB 4; Length 253;
Best Local Similarity 71.8%; Pred. No. 2.2e-32;
Matches 61; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 1 ARFNPIXRRLROINSSESPNNMTELETSMGRQSTOVQOOSKDGFINTQIRIQISAGN 60
DB 168 ARFNPIXRRLROINSSESPNNMTELETSMGRQSTOVQOOSKDGFINTQIRIQISAGN 227
QY 61 FVTXSNVRDVYSSLAIMLFECGSRP 85
DB 228 FVTXSNVRDVYSSLAIMLFECGSRP 252

RESULT 4
us-08-378-761A-77
; Sequence 77, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE

```

```

; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; TITLE OF INVENTION: USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
us-08-378-761A-77

Query Match          29.7%; Score 177; DB 1; Length 540;
Best Local Similarity 37.0%; Pred. No. 1.7e-14;
Matches 44; Conservative 18; Mismatches 53; Indels 4; Gaps 1;

QY 1 ARFNPIXRRLROINSSESPNNMTELETSMGRQSTOVQOOSKDGFINTQIRIQISAGN 60
DB 178 ARQVYIGEMKTRIRYRNRAAPDPVSITTELENSGRSLTIQESNOCATSPYQLORRNS 237
QY 61 FVTXSNVRDVYSSLAIMLFECGSRPSSLDHPSPLLRSVDAANDVTCTKSEPTVRIY 119
DB 238 KFNVDVSLIPRIALMVCAPPPSSQFS---ILIRVVNFNADVCMDEPIVRIY 292

RESULT 5
us-08-485-286-77
; Sequence 77, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```


QY 1 ARNP1XWRLRQINSSESPPMYMLETSMGROSTOVQSKDGIPTQIRLQISAGN 60
 DB 179 ARFOYLEGEMRTIRIRNRSADPPSVITLNSWGRSLTAIQSNGAFASPIQLORRNGS 238
 QY 61 FVYXSNVRDVYSSLAIMEFCSGRPPS 87
 DB 239 KFSVYDVSTILPIIALMWYRCAPPPSS 265

RESULT 8 US-07-988-430-1

; Sequence 1, Application US/07988430
 ; Patent No. 5416202
 ; GENERAL INFORMATION:
 ; APPLICANT: Bernhard, Susan L.
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Stephen F.
 ; APPLICANT: Lane, Julie A.
 ; APPLICANT: Lei, Shau-Ping
 ; TITLE OF INVENTION: Materials Comprising and Methods of
 ; PREPARATION AND USE FOR RIBOSOME-INACTIVATING PROTEINS
 ; NUMBER OF SEQUENCES: 101
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; STREET: Two First National Plaza, 20 South Clark
 ; STREET: Street
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60603
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/988,430
 ; FILING DATE: 19921209
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/901,707
 ; FILING DATE: 19-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/787,567
 ; FILING DATE: 04-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5416202and, Greta E.
 ; REGISTRATION NUMBER: 35302
 ; REFERENCE/DOCKET NUMBER: 31133
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 346-5750
 ; TELEFAX: (312) 984-9740
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-988-430-1

Query Match 23.2%; Score 138; DB 1; Length 267;
 Best Local Similarity 35.6%; Pred. No. 7.2e-10;
 Matches 31; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

QY 1 ARNP1XWRLRQINSSESPPMYMLETSMGROSTOVQSKDGIPTQIRLQISAGN 60
 DB 179 ARFOYLEGEMRTIRIRNRSADPPSVITLNSWGRSLTAIQSNGAFASPIQLORRNGS 238
 QY 61 FVYXSNVRDVYSSLAIMEFCSGRPPS 87
 DB 239 KFSVYDVSTILPIIALMWYRCAPPPSS 265

RESULT 9 US-08-218-303-16

; Sequence 16, Application US/08218303
 ; Patent No. 5547867
 ; GENERAL INFORMATION:
 ; APPLICANT: Kara, Bhupendra V.
 ; APPLICANT: Hockney, Robert C.
 ; APPLICANT: Fitton, John E.
 ; TITLE OF INVENTION: FERMENTATION PROCESS
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cushman, Darby & Cushman
 ; STREET: 1615 L Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20036-5601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/218,303
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/841,533
 ; FILING DATE: 26-FEB-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kokulis, Paul N.
 ; REGISTRATION NUMBER: 16,773
 ; REFERENCE/DOCKET NUMBER: PNK/3893/94908/MJW
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-861-3000
 ; TELEFAX: 202-822-0944
 ; TELEX: 6714627 CUSH
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-218-303-16

Query Match 23.2%; Score 138; DB 1; Length 267;
 Best Local Similarity 35.6%; Pred. No. 7.2e-10;
 Matches 31; Conservative 18; Mismatches 38; Indels 0; Gaps 0;

QY 1 ARNP1XWRLRQINSSESPPMYMLETSMGROSTOVQSKDGIPTQIRLQISAGN 60
 DB 179 ARFOYLEGEMRTIRIRNRSADPPSVITLNSWGRSLTAIQSNGAFASPIQLORRNGS 238
 QY 61 FVYXSNVRDVYSSLAIMEFCSGRPPS 87
 DB 239 KFSVYDVSTILPIIALMWYRCAPPPSS 265

RESULT 10 US-08-425-336-1

; Sequence 1, Application US/08425336
 ; Patent No. 5621083
 ; GENERAL INFORMATION:
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Stephen F.
 ; APPLICANT: Studlika, Gary M.
 ; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 ; PROTEINS
 ; NUMBER OF SEQUENCES: 140
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

RESULT 14
US-08-338-793D-61
Sequence 61, Application US/08338793D
Patent No. 5840521
GENERAL INFORMATION:
APPLICANT: Barth, Peter Thomas
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DABRY CUSHMAN
ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatibles
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338, 793D
FILING DATE: 08-No. 5840521-94
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842,081

Page 7

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Search completed: March 18, 2003, 08:28:04
Job time : 10.7739 secs


```

;
; SOFTWARE: Patentin Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,527
; FILING DATE: 18-Jan-2001
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/621,803
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27129/33199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
;
; INFORMATION FOR SEQ ID NO: 251:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-09-765-527-251

Query Match          11.8%; Score 70.5; DB 10; Length 332;
Best Local Similarity 22.6%; Pred. No. 1.1;
Matches 26; Conservative 23; Mismatches 63; Indels 3; Gaps 2;

OY 1 ARFNPXWRLRQINSGESPPNMYLETSGRSTOVQOS-KDGIFNTQIRLQISAG 59
    ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
Db 190 ARFTIENQIRN--NFOQRIRPANNITISLENKWKLSFQIRTSANGMSEAVELEBRANG 247
    : | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 60 NEVTVXSVRDYVSSLAIMLFECGSRPSSLDHSPILLSVYDAANDVCTYCTSEP 114
    : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 248 KKYVTVAVDQVPKIALKLFVDKPKRSACHHASHVAVARWASDEPFSMCAMALDP 302
    : | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
US-09-765-527-253
; Sequence 253, Application US/09765527
; Patent No. US2002000638A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; TITLE OF INVENTION: Methods for Recombinant Microbial Production of
; FUSION PROTEINS
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Tooley, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,527
; FILING DATE: 18-Jan-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/621,803
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27129/33199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 253:
```

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 253:
US-09-765-527-253

Query Match          11.7%; Score 69.5; DB 10; Length 309;
Best Local Similarity 26.4%; Pred. No. 1.3;
Matches 24; Conservative 19; Mismatches 45; Indels 3; Gaps 2;

OY 1 ARFNPXWRLRQINSGESPPNMYLETSGRSTOVQOS-KDGIFNTQIRLQISAG 59
    ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |
Db 190 ARFTIENQIRN--NFOQRIRPANNITISLENKWKLSFQIRTSANGMSEAVELEBRANG 247
    : | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 60 NEVTVXSVRDYVSSLAIMLFECGSRPSSLD 90
    : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 248 KKYVTVAVDQVPKIALKLFVDKPKRSALD 278
    : | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-738-626-6445
; Sequence 6445, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 6445
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6445

Query Match          11.7%; Score 69.5; DB 9; Length 477;
Best Local Similarity 35.4%; Pred. No. 2.3;
Matches 17; Conservative 10; Mismatches 20; Indels 1; Gaps 1;

OY 26 MLEETSGRSTOVQOSKDGIFNTQIRLQISAGNEVTVXSVRDYVSS 73
    : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 VLVVDTAHQHSTWISALK-RIRALDVNVPIVAGNVTVADGVRLVEA 287
    : | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-09-792-793A-34
; Sequence 34, Application US/09792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; APPLICANT: Coggin, Philip
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
; FILE REFERENCE: 25020-601D
```

```

; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 34
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Bryonia dioica
US-09-792-793A-34

Query Match      11.6%; Score 69; DB 9; Length 247;
Best Local Similarity 25.3%; Pred. No. 1.1;
Matches 21; Conservative 25; Mismatches 29; Indels 8; Gaps 4;

OY 1 ARFNPIKRLRRQINSSESPNNMYMLETSGRSTQVO--OSKDGIFNTQIRLOISA 58
Db 162 ARKRFIEQIGKRVN--KTFPLSLATISLENNMSALSKQIOIASNNQGFSPVVL-IDG 218
OY 59 GN---FVYXSNVRYSSLAIML 78
Db 219 NNQRVSTNNSARVYTSNIALLL 241

RESULT 9
US-09-765-527-247
; Sequence 247, Application US/09765527
; Patent No. US20020006638A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; TITLE OF INVENTION: Methods for Recombinant Microbial Production of
; Fusion Proteins and BRI-Derived Peptides
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,527
; FILING DATE: 18-Jan-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/621,803
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27129/33199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 247:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 247:
US-09-765-527-247

Query Match      11.5%; Score 68.5; DB 10; Length 251;
Best Local Similarity 27.5%; Pred. No. 1.3;
Matches 22; Conservative 17; Mismatches 38; Indels 3; Gaps 2;

OY 1 ARFNPIKRLRRQINSSESPNNMYMLETSGRSTQVOQS-KDGIFNTQIRLOISAG 59
||| |::| | : | : || |::| :|::| :|::|
```

```

Db 168 ARFFLENIQIRN--NFQQRIRPANNITISLENNKGLSFQIRISGANGMSEAVELEERANG 225
OY 60 NEVYXSNVRYSSIAIMLF 79
Db 226 KKYVTVAVDQVKPKIALKRF 245

RESULT 10
US-09-801-574-32
; Sequence 32, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Peijng Jeremy
; APPLICANT: Page, David C.
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399,2007-002
; CURRENT APPLICATION NUMBER: US/09/801,574
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 1018
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-801-574-32

Query Match      11.5%; Score 68.5; DB 10; Length 1018;
Best Local Similarity 34.8%; Pred. No. 8.2;
Matches 23; Conservative 10; Mismatches 22; Indels 11; Gaps 3;

OY 9 RLRRQINSSESPNNMYMLET-----SMGRSTQVOQSK-DGIFNTQIRLOI-----S 57
Db 213 RLRSIGDACEAKRPDENQALRLSKNPIPLMGRRSNQSTFSRPKGHFNRLILDVKEYTS 272
OY 58 AGNFVPT 63
Db 273 AGNHVT 278

RESULT 11
US-09-949-192-19
; Sequence 19, Application US/09949192
; Patent No. US20020142292A1
; GENERAL INFORMATION:
; APPLICANT: Parham, Christi L.
; APPLICANT: Gorman, Daniel L.
; APPLICANT: Kurata, Hirokazu
; APPLICANT: Arai, Naoko
; APPLICANT: Sana, Theodore R.
; APPLICANT: Mattson, Jeanine D.
; APPLICANT: Murphy, Erin E.
; APPLICANT: Savkoor, Chetan
; APPLICANT: Grein, Jeffrey
; APPLICANT: Smith, Kathleen M.
; APPLICANT: McClanahan, Terill K.
; TITLE OF INVENTION: MAMMALIAN GENES, RELATED REAGENTS AND METHODS
; FILE REFERENCE: DX01169K
; CURRENT APPLICATION NUMBER: US/09/949,192
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/231,267
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 19
; LENGTH: 937
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-949-192-19
```



```

: PRIOR APPLICATION NUMBER: JP00/00966
: PRIOR FILING DATE: 2000-02-21
: PRIOR APPLICATION NUMBER: 11-44890
: PRIOR FILING DATE: 1999-02-23
: PRIOR APPLICATION NUMBER: 11-198852
: PRIOR FILING DATE: 1999-07-13
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 15
: LENGTH: 951
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Amino acid sequence of agarase 4-3
US-09-924-097-15

Query Match                      10.8%; Score 64.5; DB 9; Length 951;
Best Local Similarity            24.8%; Pred. No. 23;
Matches 27; Conservative 18; Mismatches 51; Indels 13; Gaps 5;

QY      15 NSGSEPPMMYML--ELETSNGRSTQYQOSKDDIFMTQIRLQISAGNFYTSNVDVYS 72
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      48 NDGSSPIIITYYNGEQALISWYNAGDYDVYVAV--GGEYNIIELVQTSVAANPIIEVLV 106

QY      73 SLAIMFECSGR---PFSSLHPSFLILRSYVDAANDYTCTXSEPTVRI 118
          |||---|---|---|---|---|---|---|---|---|---|---|---|
Db      107 D-SNGTFEESQGYIVPLGSMEDFQPLV-----AAHTVTLPAQTSIRL 148

RESULT 14
US-09-347-064-10
: Sequence 10, Application US/09347064A
: Patent No. US20020045208A1
: GENERAL INFORMATION:
: APPLICANT: Eck, Jurgen
: APPLICANT: Schmidt, Arno
: APPLICANT: Zinke, Holger
: TITLE OF INVENTION: Recombinant Fusion Proteins Based on
: TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
: FILE REFERENCE: 09282-5
: CURRENT APPLICATION NUMBER: US/09/347,064A
: EARLIER FILING DATE: 1999-07-02
: EARLIER APPLICATION NUMBER: PCT/EP98/00009
: EARLIER FILING DATE: 1998-01-02
: EARLIER APPLICATION NUMBER: EP 97 10 0012.0
: EARLIER FILING DATE: 1997-01-02
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 10
: LENGTH: 263
: TYPE: PRT
: ORGANISM: Viscum album
US-09-347-064-10

Query Match                      10.7%; Score 64; DB 10; Length 263;
Best Local Similarity            80.0%; Pred. No. 4.8;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

QY      105 NDVYCTXSEPTVRIV 119
          :||||:|||||
Db      1 DDVTCASASEPTVRIV 15

RESULT 15
US-09-347-064-4
: Sequence 4, Application US/09347064A
: Patent No. US20020045208A1
: GENERAL INFORMATION:
: APPLICANT: Eck, Jurgen
: APPLICANT: Schmidt, Arno
: APPLICANT: Zinke, Holger
: TITLE OF INVENTION: Recombinant Fusion Proteins Based on

```

; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-4

Query Match 10.7%; Score 64; DB 10; Length 267;
Best Local Similarity 80.0%; Pred. No. 4.9;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 105 NDVCTXSEPTVRIY 119
:||||:|||||||
Db 1 DDVTCASSEPTVRIY 15

Search completed: March 18, 2003, 08:28:37
Job time : 9.57789 secs

RESULT 3

RICOSAG
agglutinin precursor - castor bean
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Ricinus communis (castor bean)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
A:Accession: A24261; A24210
R:Roberts, L.M.; Lamb, I.F.; Pappin, D.J.C.; Lord, J.M.
J.R. Biol. Chem. 260, 15682-15685, 1985
A>Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.
A:Reference number: A24261; MUID:86059449; PMID:299130

A:Molecule type: mRNA
A:Residues: 1-564 <ROB>
A:Cross-references: GB:I12089; NID:g169700; PIDN:AAA33869.1; PID:g169701
R:Araiki, T.; Yoshikawa, Y.; Funatsu, G.
Biochim. Biophys. Acta 872, 277-285, 1986
A>Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.
A:Reference number: A24210
A:Accession: A24210

A:Molecule type: protein
A:Residues: 303-325,'F','327-330','T','332-361','D','363-373','G','375-403','T','405-551','V','S'
C:Comment: This protein has strong agglutinating activity and weak cytotoxicity comparsed
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; s
E:1.24/Domains: signal sequence #status predicted<SIG>
F:125-220/Product: agglutinin chain A #status predicted<ACH>
F:303-281/Product: agglutinin chain B #status predicted<BGCH>
F:303-564/Product: agglutinin chain B #status experimental<BGCH>
F:319-361,362-402,405-443,450-485,489-528,531-564/Region: 40-residue repeats
F:34,259/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:104,147,231,233/Binding site: substrate (Thr,Tyr,Glu,Asn) #status predicted
F:200,203/Active site: Glu,Arg #status predicted
F:282-306,322-341,365-382,453-466,492-509/Dissulfide bonds: #status predicted
F:337,343,348/Binding site: carbohydrate (Asp,Asn) #status predicted
F:334,437/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:536,557/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 29.7% Score 177; DB 1; Length 564;
Best local Similarity 37.0%; Pred. No. 7.5e+12;

Matches 44; Conservative 18; Mismatches 53; Indels 4; Gaps 1;

OY 1 ARPNPLXWRRLQINSGESSPPNNMYMLEETSNKGSQTVOVGSKDIGINTIRIQISAGN 60
Db 202 ARFOYLEGMRTIRIRNRFSADDPSTITLWNSWGRLTAIOENSGAFASPIQLQRNGS 261
OY 61 FVTYSNVDRVISLAIMLEFGCGRPSSLDHRSPLLARSVDAAANDVCIXSEPTARI V 119
Db 262 KENVIVDSILPILLIWMYYRCAPPSQS----LLTRPVVPNMADVCDMDPEPIRV IV 316

RESULT 4

TZLSA
ablin-a precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian Licorice)
C>Date: 31-Dec-1993 #sequence_revision 01-Aug-1997 #text_change 16-Jun-1999
A:Accession: S32429; JT002; A39761; JC1398; SI4473; S24133; S74111
R:Huang, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A>Title: Primary structure of three distinct isoforms determined by cDNA sequencing.
A:Reference number: S32429; MUID:93132798; PMID:8421313
A:Accession: S32429
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'E','-2-528'<HDN>
A:Cross-references: GB:998344; NID:g166294; PITN:AAA23624.1; DTG:d16670E

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
 C/Accession: S16022
 R/Mood: K.A.: Lord, J.M.: Mawzynczak, E.J.; Platak, M.
 Eur. J. Biochem. 198, 723-732, 1991
 A>Title: Preprobrin: genomic cloning, characterisation and the expression of the A-chain
 A/Reference number: S16022; MUID:91266957; PMID:2050149
 A/Accession: S16022
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-562 <MOO>
 A:Cross-references: EMBL:X55667; NID:g16084; PIDN:CAA39202.1; PID:g16085
 C:Comment: Brin consists of an A chain, which inhibits protein synthesis by inactivating the A and B chains are linked by a single disulfide bond, which is essential for toxicity
 C:Superfamily: ricin; RNA N-glycosidase homology
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglytamic acid
 F:35-285/Product: abrin-c chain A #status predicted <ACH>
 F:41-280/Domain: rRNA N-glycosidase homology <RNG>
 F:295-562/Product: abrin-c chain B #status predicted <BCH>
 F:317-359,360-400,403-441,448-483,487-526,529-562/Region: 40-residue repeats
 F:35/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predicted
 F:108,147,229,230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:198,201/Active site: Glu, Arg #status predicted
 F:234,287,395,435,436/Binding site: carbonyl (Asn) (covalent) #status predicted
 F:281-303,320-339,363-380,451-464,490-507/Disulfide bonds: #status predicted
 F:322,346/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
 F:534,555/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 23.2%; Score 138; DB 2; Length 562;
 Best Local Similarity 36.1%; Pred. No. 1.9e-07;
 Matches 43; Conservative 20; Mismatches 50; Indels 6; Gaps 4;

OY 1 ARFNPIXRRLRQINSSESPNNMMLLETSMGROSTOVQOSKDGIFNTQIRLOISAG 59
 DB 200 ARRYRISNVGVISIRTFAPQDPAAMLSLENMNDLSGVQOVSQVDFPPNVAITLRSNQ 259
 OY 60 NFVTSXNVADV--TSSLAIMLEFCGSRPFSSLDHPSPLLRSVYDANDVCTCXSEPTVRI 118
 DB 260 PVTVDLSLHPTVAVALMLFVCN--PPNA--NQSPILRSIVEERK--ICSSRYEPTVRI 313

RESULT 7
 S32430
 abrin-b precursor - Indian licorice (fragment)
 N/Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C/Species: Abrus precatorius (Indian licorice)
 C/Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 20-Aug-1999
 C/Accession: S32430; JCI399
 R/Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
 J. Mol. Biol. 229, 263-267, 1993
 A>Title: Primary structure of three distinct isoabrin determined by cDNA sequencing. CD
 A/Reference number: S32429; MUID:93132798; PMID:8421313
 A/Accession: S32430
 A:Molecule type: mRNA
 A:Residues: 1-527 <HON>
 A:Cross-references: GB:M98345; NID:g166296; PIDN:AAA32625.1; PID:g166297
 R/Kimura, M.; Sumizawa, T.; Funatsu, G.
 Biosci. Biotechnol. Biochem. 57, 166-169, 1993
 A>Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxin
 A/Reference number: JCI398; MUID:93169023; PMID:7763422
 A/Accession: JCI399
 A:Molecule type: protein
 A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'M', 427, 'D', 429-430
 C:Experimental source: seed
 C:Superfamily: ricin; RNA N-glycosidase homology
 C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; F
 F:1-250/Product: abrin-b chain A #status predicted <ACH>
 F:7-245/Domain: rRNA N-glycosidase homology <RNG>
 F:260-527/Product: abrin-b chain B #status experimental <BCH>
 F:282-324,325-365,368-406,413-448,452-491,494-527/Region: 40-residue repeats
 F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status predicted
 F:74,113,194,195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:110,360,400/Binding site: carbonyl (Asn) (covalent) #status predicted
 F:163,166/Active site: Glu, Arg #status predicted

F:246-268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted
 F:287,311/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
 F:499,520/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 22.7%; Score 135; DB 2; Length 527;
 Best Local Similarity 33.3%; Pred. No. 3.9e-07;
 Matches 40; Conservative 24; Mismatches 48; Indels 8; Gaps 5;

OY 1 ARFNPIXRRLRQINSSESPNNMMLLETSMGROSTOVQOSKDGIFNTQIRLOISAGN 60
 DB 165 ARRYRISNVGVISIRTFAPQDPAAMLSLENMNDLSGVQOVSQVDFPPNVAITLRSVNN 223
 OY 61 --FVTSXNVADV--TSSLAIMLEFCGSRPFSSLDHPSPLLRSVYDANDVCTCXSEPTVRI 118
 DB 224 QPVTVDLSLHPTVAVALMLFVCN--PPNA--NQSPILRSIVEERK--ICSSRYEPTVRI 278

RESULT 8
 S62627
 agglutinin I precursor - European elder
 C/Species: Sambucus nigra (European elder)
 C/Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 20-Aug-1999
 C/Accession: S62627; S62619
 R/van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.
 Eur. J. Biochem. 235, 128-137, 1996
 A>Title: The NeutG(alpha-2,6)-Gal/GalNAc-binding lectin from elderberry (Sambucus nigra)
 A/Reference number: S62619; MUID:96202926; PMID:8631319
 A/Accession: S62627
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-570 <VAN>
 A:Cross-references: EMBL:U27122; NID:g141772; PIDN:AAC49158.1; PID:g141773
 A:Accession: S62619
 A:Molecule type: protein
 A:Residues: 29-39;309-319 <VA2>
 C:Superfamily: ricin; RNA N-glycosidase homology
 F:37-283/Domain: rRNA N-glycosidase homology <RNG>

Query Match 20.4%; Score 121.5; DB 2; Length 570;
 Best Local Similarity 25.2%; Pred. No. 1.4e-05;
 Matches 32; Conservative 31; Mismatches 53; Indels 11; Gaps 4;

OY 1 ARFNPIXRRLRQINSSESPNNMMLLETSMGROSTOVQOSK--DGIRNTQIRLOISAG 59
 DB 201 ARRYRIELRIRTSITDASERTPDLLMSNMNMSMSSETQDAQPGCIGAGVQLRDERN 260
 OY 60 NFVTSXNVADV--TSSLAIMLEFCGSRPFSSLDHPSPLLRSVY-----DANDVCTCX 111
 DB 261 NSIEVNFRLFEITFYIAVLVYCA--PVTSSYSNNAIDAOIKMPVFRGGEYEVCSV 318
 OY 112 SEPTVRI 118
 DB 319 VEYTRRI 325

RESULT 9
 B39761
 abrin (clone 3.7) precursor - Indian licorice (fragment)
 N/Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C/Species: Abrus precatorius (Indian licorice)
 C/Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 05-Jun-1998
 C/Accession: B39761
 R/Evensen, G.; Mathiesen, A.; Sundan, A.
 J. Biol. Chem. 266, 6848-6852, 1991
 A>Title: Direct molecular cloning and expression of two distinct abrin A-chains.
 A/Reference number: A39761; MUID:91201329; PMID:2016300
 A/Accession: B39761
 A:Molecule type: DNA
 A:Residues: 1-106 <EVE>
 A:Cross-references: GB:X54873
 C:Superfamily: ricin; RNA N-glycosidase homology
 C:Keywords: duplication; glycosidase; hydrolase; lectin; toxin

A:Molecule type: DNA
 A:Residues: 1-294 <KAT>
 A:Cross-references: EMBL:D10600; NID:g218010; PIDN:BAA01451.1; PID:g218011
 C:Superfamily: rRNA N-glycosidase
 C:Keywords: glycosidase; hydrolase
 F:30-278/Domain: rRNA N-glycosidase homology <RNG>

Query Match 14.1%; Score 84; DB 2; Length 294;
 Best Local Similarity 26.6%; Pred. No. 0.11;
 Matches 21; Conservative 19; Mismatches 37; Indels 2; Gaps 1;

OY 1 ARFNPXWRLRQINSSESPNNMYLETSMGROSTOVQOSKDGIFNTOIRLOISA 60
 DB 201 ARFKTIENQVKTNN--RAFPNARVNLIESWGKISTAIHNAKNGALTSPLKLNANGS 258
 OY 61 EFTXSNRVDVTSILMIF 79
 DB 259 KWIVLRVDIDPVDLLK 277

RESULT 13
 JE0401
 antiviral protein - Virginian pokeweed
 C:Species: Phytolacca americana (Virginian pokeweed)
 C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 07-May-1999
 C:Accession: JE0401
 R:Kung, S.S.; Kimura, M.; Funatsu, G.
 Agric. Biol. Chem. 54, 3301-3318, 1990
 A:Title: The complete amino acid sequence of antiviral protein from the seeds of pokeweed
 A:Reference number: JE0401; MUID:91242096; PMID:1368643
 A:Accession: JE0401
 A:Molecule type: protein
 A:Residues: 1-261 <KUN>
 A:Experimental source: seed
 C:Comment: This protein prevents the replication of a number of plant viruses, and inhibits
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
 C:Keywords: disulfide bond; glycoprotein
 F:6-254/Domain: rRNA N-glycosidase homology <RNG>
 F:10/44,255/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:34-258,84-105/Disulfide bonds: #status experimental

Query Match 13.3%; Score 79; DB 2; Length 261;
 Best Local Similarity 33.9%; Pred. No. 0.34;
 Matches 20; Conservative 13; Mismatches 25; Indels 2; Gaps 1;

OY 1 ARFNPXWRLRQINSSESPNNMYLETSMGROSTOVQOSKDGIFNTOIRLOISA 59
 DB 177 ARFKTIENQVKTNNRDFS--PNDKVLDEENMGKISTAIHNSKNGALPRLKLNADG 233

RESULT 14
 JC5032
 karasurin-B - Trichosanthes kirilowii var. japonica
 C:Species: Trichosanthes kirilowii var. japonica
 C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 23-May-1997
 C:Accession: JC5032
 R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.
 Biol. Pharm. Bull. 19, 1485-1489, 1996
 A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and karasurin-A
 A:Reference number: JC5032; MUID:97108848; PMID:8951169
 A:Accession: JC5032
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1247 <KON>
 C:Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyze
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
 F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 13.1%; Score 78; DB 2; Length 247;
 Best Local Similarity 28.0%; Pred. No. 0.41;
 Matches 23; Conservative 23; Mismatches 30; Indels 6; Gaps 3;

OY 1 ARFNPXWRLRQINSSESPNNMYLETSMGROSTOVQ--QSKDGFNTOIRLOISA 58

DB 162 ARFKTIENQVKTNN--KTFPLSLATISLENSALSLSKQIOIASTNNGOFEITPVVLINAO 219
 OY 59 GNFVYXSNVRD--VITSLAIML 78
 DB 220 NORVITITNVAGVVTSMIALLL---NRNMAAIDDDVPM 278

RESULT 15
 JC5606
 karasurin C - Trichosanthes kirilowii var. japonica
 N:Contains: karasurin A
 C:Species: Trichosanthes kirilowii var. japonica
 C:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 19-Jul-2002
 C:Accession: JC5606; JC5033
 R:Mizukami, H.; Iida, K.; Kondo, T.; Ogihara, Y.
 Biol. Pharm. Bull. 20, 711-713, 1997
 A:Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating protein
 A:Reference number: JC5606; MUID:97356562; PMID:9212998
 A:Accession: JC5606

A:Molecule type: DNA
 A:Residues: 1-289 <MTZ>
 A:Cross-references: DDBJ:AB000666; NID:g2329830; PIDN:BAA21786.1; PID:g2329831
 R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.
 Biol. Pharm. Bull. 19, 1485-1489, 1996
 A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and karasurin-A
 A:Reference number: JC5032; MUID:97108848; PMID:8951169
 A:Accession: JC5033
 A:Molecule type: protein
 A:Residues: 22-270 <KON>
 C:Comment: This protein is a ribosome-inactivating protein and exhibits cytotoxic, abortifacient, and
 C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
 F:22-270/Product: karasurin C #status predicted <MAC>
 F:24-270/Product: karasurin A #status predicted <MAC>
 F:27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 12.8%; Score 76.5; DB 2; Length 289;
 Best Local Similarity 25.3%; Pred. No. 0.74;
 Matches 25; Conservative 28; Mismatches 37; Indels 9; Gaps 4;

OY 1 ARFNPXWRLRQINSSESPNNMYLETSMGROSTOVQ--QSKDGFNTOIRLOISA 58
 DB 185 ARFKTIENQVKTNN--KTFPLSLATISLENSALSLSKQIOIASTNNGOFEITPVVLINAO 242
 OY 59 GNFVYXSNVRD--VITSLAIMLFECSGRPFSSLDHPSPL 95
 DB 243 NORVITITNVAGVVTSMIALLL---NRNMAAIDDDVPM 278

Search completed: March 18, 2003, 08:27:30
 Job time : 15.8668 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:01:23 ; Search time 5.38191 Seconds
(without alignments)
917.088 Million cell updates/sec

Title: US-09-627-165E-14

Perfect score: 596
Sequence: 1 ARNPPIKWRMRQINGESS.....VVDANDVTCTXSEPTVRIV 119

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwisProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	317	53.2	254	1	MLA_VISAL
2	178	29.9	576	1	P02879 rictinus com
3	177	29.7	564	1	AGGL_RICCO
4	148	24.8	528	1	ABRA_ABRPR
5	138	23.2	562	1	ABRR_ABRPR
6	135	22.7	527	1	ABRB_ABRPR
7	108	18.1	563	1	NIGB_SAMNI
8	92	15.4	286	1	RIP1_MOMCH
9	84	14.1	284	1	RIPA_PHYAM
10	83.5	14.0	282	1	RIP2_BRIDI
11	79	13.3	261	1	RIPS_PHYAM
12	76.5	12.8	289	1	RIPS_TRIKI
13	74	12.8	754	1	KATC_ARATH
14	74	12.4	277	1	RIPA_LUCFY
15	74	12.4	286	1	RIP1_CUCFI
16	74	12.4	313	1	RIP1_PHYAM
17	74	12.4	730	1	EF2_MERPU
18	72.5	12.2	289	1	RIP1_TRIKI
19	70	11.7	730	1	EF2_METMT
20	69.5	11.7	1054	1	RDGB_DROME
21	69	11.6	290	1	RIP1_BRYDI
22	68.5	11.5	268	1	CAPA_YEAST
23	68.5	11.5	316	1	RIPG_GELMU
24	68.5	11.5	1018	1	SR31_MOUSE
25	67.5	11.3	230	1	RIPB_LUCFY
26	67	11.2	294	1	RIP1_TRIAN
27	67	11.2	572	1	EF2_HALHA
28	66	11.1	728	1	EF2_MERTE
29	65.5	11.0	372	1	PLGI_BUCAI
30	65.5	11.0	966	1	Y918_HUMAN
31	65.5	11.0	1233	1	IR51_MOUSE
32	65.5	11.0	1267	1	V13_REOVL
33	65.5	11.0	1267	1	V13_REOVL

34	65	10.9	977	1	YD68_SCHPO
35	65	10.9	1211	1	PKP4_HUMAN
36	64.5	10.8	389	1	YND5_YEAST
37	64.5	10.8	435	1	PNCB_VIBCH
38	64.5	10.8	518	1	COPD_ORYSA
39	64	10.7	601	1	MUTL_LISMO
40	64	10.7	730	1	EF2_MERTE
41	64	10.7	836	1	PDG2_CANAL
42	63.5	10.7	603	1	MUTL_LISIN
43	63.5	10.7	742	1	KM65_YEAST
44	63	10.6	317	1	CAH6_MOUSE
45	63	10.6	872	1	DP3A_MYCPN

ALIGNMENTS

RESULT 1	MLA_VISAL	STANDARD:	PRT:	254 AA.
AC	P81446:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (IRNA N-glycosidase) (EC 3.2.2.22).			
OS	Viscum album (European mistletoe).			
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Santalales; Viscaceae; Viscum.			
OX	NCBI_Taxid=3972;			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN-Subsp. album;			
RX	MEDLINE=97134581; PubMed=8980141;			
RA	Huquet Soler M., Scoeva S., Schwaborn C., Wilhelm S., Stiefel T.,			
RA	Voelter W.;			
RT	"Complete amino acid sequence of the A chain of mistletoe lectin I.";			
FEBS Lett. 399:153-157(1996).				
CC	-I- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY INTO THE CELL OF THE A CHAIN. B CHAINS ARE ALSO RESPONSIBLE FOR CELL AGGLUTINATION (LECTIN ACTIVITY).			
CC	-I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.			
CC	-I- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.			
CC	-I- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA' AND NON-GLYCOSYLATED FORM MLA''.			
CC	-I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.			
CC	HSSP: P11140; IABR.			
DR	Interpro: IPR001574; RIP.			
DR	PRINTS: PR000396; SHIGARICIN.			
DR	PROSITE: PS00275; SHIGA_RICIN; Repeat: Glycoprotein; Lectin.			
KW	Plant defense; Hydrolase; Toxin; Repeat: Glycoprotein; Lectin.			
FT	ACET SITE	165	112	165
FT	CARBOHYD	112	112	112
FT	VAR	15	15	15
FT	VAR	66	66	66
FT	VAR	112	112	112
FT	VAR	116	116	116
FT	VAR	133	133	133
FT	VAR	140	140	140
FT	VAR	144	144	144
FT	VAR	151	151	151
FT	VAR	179	179	179
FT	VAR	184	184	184
FT	VAR	190	190	190
FT	VAR	218	218	218
FT	VAR	223	223	223
FT	VAR	224	224	224
FT	VAR	223	223	223

```

FT VARIANT 231 231 T -> S (IN M.L.A.).
FT VARIANT 235 235 D -> S (IN M.L.A.).
SQ SEQUENCE 254 AA; 28478 MW; 53ABF98D3E0FEE67 CRC64;
Query Match 53 2%; Score 317; DB 1; Length 254;
Best Local Similarity 71.6%; Pred.No. 3.7e-29;
Matches 63; Conservative 7; Mismatches 18; Indels 0; Gaps 0;
OY 1 AAFNEIKWRLRQINSGESSPPNMYMLELFTSGRQSTQVOQSKDGFNTQIRLQISAGN 60
DB 167 AAFNFIIMRYROXINSGASFLPDVYMLELFTSGQOSTQOVHSTDGVEFNPIRLAIPGN 226
OY 61 FVTXSNVADVTSIAIMLFECSSGRPFSS 88
DB 227 FVTLTINVDVIAIAIMLFVCGGRPFSS 254

RESULT 2
R1CL_RICCO STANDARD; PRT; 576 AA.
AC P02879; P02880;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ricin precursor [contains: Ricin A chain (rRNA N-glycosidase)
OS (EC 3.2.2.22); Ricin B chain].
OC Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Ericosids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxId=3988;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=86067214; PubMed=2999712;
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,
RA Weaver R.F.;
RT "Genomic cloning and characterization of a ricin gene from Ricinus
RT communis.";
RL Nucleic Acids Res. 13:8019-8033(1985).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=92163016; PubMed=1371405;
RA Tregear J.W., Roberts L.M.;
RT "The lectin gene family of Ricinus communis: cloning of a functional
RT ricin gene and three lectin pseudogenes.";
RL Plant Mol. Biol. 18:515-525(1992).
RN (3)
RP SEQUENCE OF 12-576 FROM N.A.;
RX MEDLINE=85179479; PubMed=3838723;
RA Lamb A., Roberts L.M., Lord J.M.;
RT "Nucleotide sequence of cloned cDNA coding for preproricin.";
RL Eur. J. Biochem. 148:265-270(1985).
RN (4)
RP SEQUENCE OF 36-302.
RA Yoshitake S., Funatsu G., Funatsu M.;
RT "Isolation and sequences of peptic peptides, and the complete
RT sequence of Ile chain of ricin-D.";
RL Agric. Biol. Chem. 42:1267-1274(1978).
RN (5)
RP SEQUENCE OF 315-576.
RA Funatsu G., Kimura M., Funatsu M.;
RT "Primary structure of Ala chain of ricin D.";
RL Agric. Biol. Chem. 43:2221-2224(1979).
RN (6)
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
RX MEDLINE=90344223; PubMed=1368517;
RA Kimura Y., Kusunoku H., Tada M., Takagi S., Funatsu G.;
RT "Structural analyses of sugar chains from ricin A-chain variant.";
RL Agric. Biol. Chem. 54:157-162(1990).
RN (7)
RP REVIEW.
RX MEDLINE=21480122; PubMed=11595634;
RA Olmes S., Kozlov J.V.;

```

```

RT "Ricin.";
RL Toxicon 39:1723-1728(2001).
RN (8)
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=87165983; PubMed=3558397;
RA Monfort W., Villafraña J.E., Monzinger A.F., Ernst S.R., Katzin B.,
RA Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;
RT "The three-dimensional structure of ricin at 2.8 A.";
RL J. Biol. Chem. 262:5398-5403(1987).
RN (9)
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
RX MEDLINE=91352004; PubMed=1881881;
RA Katzin B.J., Collins E.J., Robertus J.D.;
RT "Structure of ricin A-chain at 2.5 A.";
RL Proteins 10:251-259(1991).
RN (10)
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.
RX MEDLINE=91352005; PubMed=1881882;
RA Rutenber E., Robertus J.D.;
RT "Structure of ricin B-chain at 2.5-A resolution.";
RL Proteins 10:260-269(1991).
RN (11)
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.
RX MEDLINE=95082010; PubMed=7990130;
RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,
RA Pauprit R.A.;
RT "X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";
RL J. Mol. Biol. 244:410-422(1994).
RN (12)
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.
RX MEDLINE=96574422; PubMed=8780513;
RA Day P.J., Ernst S.R., Frankel A.E., Monzinger A.F., Pascal J.M.,
RA Molina-Svinth M.C., Robertus J.D.;
RT "Structure and activity of an active site substitution of ricin A
RT chain.";
RL Biochemistry 35:11098-11103(1996).
RN (13)
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
RX MEDLINE=97240820; PubMed=9086280;
RA Yan X., Hollis T., Svinth M., Day P., Monzinger A.F., Milne G.W.,
RA Robertus J.D.;
RT "Structure-based identification of a ricin inhibitor.";
RL J. Mol. Biol. 266:1043-1049(1997).
RN (14)
RP MOTAGENESIS.
RX MEDLINE=93165632; PubMed=1287657;
RA Kin Y., Robertus J.D.;
RT "Analysis of several key active site residues of ricin A chain by
RT mutagenesis and X-ray crystallography.";
RL Protein Eng. 5:775-779(1992).
CC -I- FUNCTION: Ricin is highly toxic to animal cells and to a less
CC extent to plant cells. The A chain is responsible for inhibiting
CC protein synthesis through the catalytic inactivation of 60S
CC ribosomal subunits. It acts as a glycosylase that removes a
CC specific adenine residue from an exposed loop of 28S ribosomal
CC RNA. As this loop is involved in the binding of elongation
CC factors, the modified ribosomes are unable to support protein
CC synthesis. The A chain can inactivate a few thousand ribosomes
CC per minute, thus inactivating them faster than the cell can make
CC new ones. A single A-chain molecule can therefore kill an animal
CC cell. The B chain binds to cell receptors and facilitates the
CC entry into the cell of the A chain. B chains are also responsible
CC for cell agglutination (lectin activity). It binds to beta-D-
CC galactopyranoside moieties.
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -I- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -I- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -I- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE
CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271.
CC -I- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.

```


DR SMART: SM00458; RICIN: 2.
 DR PROSITE: PS50231; RICIN_B-LECTIN; 2.
 DR PROSITE: PS00275; SHIGA-RICIN; 1.
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KW Glycoprotein; Lectin; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 290
 FT PROPEP 291 302
 FT CHAIN 303 564
 FT DOMAIN 309 436
 FT DOMAIN 439 563
 FT REPEAT 319 361
 FT REPEAT 362 402
 FT REPEAT 405 437
 FT REPEAT 450 485
 FT REPEAT 489 528
 FT REPEAT 531 558
 FT ACT_SITE 200 200
 FT DISULFID 282 306
 FT DISULFID 322 341
 FT DISULFID 365 382
 FT DISULFID 453 466
 FT DISULFID 492 509
 FT CARBOHYD 34 34
 FT CARBOHYD 259 259
 FT CARBOHYD 397 397
 FT CARBOHYD 437 437
 FT CARBOHYD 331 331
 FT CONFLICT 362 362
 FT CONFLICT 374 374
 FT CONFLICT 404 404
 FT CONFLICT 552 552
 SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CMC64;
 Query Match 29.7%; Score 177; DB 1; Length 564;
 Best Local Similarity 37.0%; Pred. No. 1.4e-12;
 Matches 44; Conservative 18; Mismatches 53; Indels 4; Gaps 1;
 QY 1 ARNPXMRIRROINSSESPRMVYLETSMGRSGTQVOGSKDGFNTQIRLOISAGN 60
 DB 202 ARFOYLEGEMRTIRNRNSAPPSVITLNSWGRSLTALQSNQGAFAFSPQLORRNGS 261
 QY 61 FVFXSNRVDISSLAIFECSSGRPSYSLDHPDLLRSVDAANDVCTGXSEPTVRIV 119
 DB 262 KENVYDVSIPIITIALMVYNCAPPPSSQFS---LLIRVVNPNFNDVCMDEPIVRIV 316
 RESULT 4
 ABRA_ABRPR STANDARD; PRT; 528 AA.
 ID ABRA_ABRPR
 AC P11140; P28589;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Abrin-a precursor [Contains: Abrin-a A chain (rRNA N-glycosidase)
 (EC 3.2.2.22); Abrin-a B chain].
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 OC NCBI_TaxID=3816;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93132798; PubMed=8421313;
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
 RT "Primary structure of three distinct isoabrinins determined by cDNA
 sequencing. Conservation and significance.";
 RL J. Mol. Biol. 229:263-267(1993).
 RN [2]
 RN SEQUENCE OF 1-251.
 RP TISSUE=Seed;
 RC Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
 RA "The complete amino acid sequence of the A-chain of abrin-a, a toxic

RT protein from the seeds of Abrus precatorius.";
 RL Agric. Biol. Chem. 52:1095-1097(1988).
 RN [3]
 RN SEQUENCE OF 1-251 FROM N.A.
 RP TISSUE=Leaf;
 RC MEDLINE=91201329; PubMed=2016300;
 RX Evensen G., Mathiesen A., Sundan A.;
 RA "Direct molecular cloning and expression of two distinct abrin
 A-chains.";
 RL J. Biol. Chem. 266:6848-6852(1991).
 RN [4]
 RN SEQUENCE OF 262-528.
 RP MEDLINE=92371656; PubMed=1505674;
 RX Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
 RA "The complete primary structure of abrin-a B chain.";
 RL FEBS Lett. 309:115-118(1992).
 RN [5]
 RN X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
 RP MEDLINE=95333188; PubMed=7608980;
 RX Tahlrov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
 RA "Crystal structure of abrin-a at 2.14 A.";
 RL J. Mol. Biol. 250:354-367(1995).
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
 FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
 PRECEDES ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
 INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M9344; AAA32624.1; ALT_INT.
 DR EMBL: X54872; -; NOT_ANNOTATED_CDS.
 DR PIR: S2429; T2LSA.
 DR PIR: S24133; S24133.
 DR PDB: 1ABR; 07-FEB-95.
 DR InterPro: IPR001574; RIP.
 DR InterPro: IPR000772; Ricin_B-lectin.
 DR Pfam: PF00161; RIP; 1.
 DR Pfam: PF00652; Ricin_B-lectin; 6.
 DR PRINTS: PR00396; SHIGARICIN.
 DR SMART: SM00458; RICIN; 2.
 DR PROSITE: PS50231; RICIN_B-LECTIN; 2.
 DR PROSITE: PS00275; SHIGA-RICIN; 1.
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KW Glycoprotein; Lectin; 3D-structure.
 FT CHAIN 1 251
 FT PEPTIDE 252 261
 FT CHAIN 262 528
 FT DOMAIN 273 400
 FT DOMAIN 403 527
 FT REPEAT 283 325
 FT REPEAT 326 366
 FT REPEAT 369 401
 FT REPEAT 414 449
 FT REPEAT 453 492
 FT REPEAT 495 528
 FT ACT_SITE 164 164
 BY SIMILARITY.

```

FT DISULFID 247 269 INTERCHAIN (BY SIMILARITY).
FT DISULFID 286 305 BY SIMILARITY.
FT DISULFID 329 346 BY SIMILARITY.
FT DISULFID 417 430 BY SIMILARITY.
FT DISULFID 456 473 BY SIMILARITY.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).
FT CONFLICT 401 401 N-LINKED (GLCNAC. . .).
FT CONFLICT 202 202 MISSING (IN REF. 2).
FT CONFLICT 298 298 M -> Y (IN REF. 4).
FT CONFLICT 427 427 M -> L (IN REF. 4).
FT CONFLICT 467 467 T -> P (IN REF. 4).
FT CONFLICT 483 483 V -> L (IN REF. 4).
SQ SEQUENCE 528 AA; 59243 MW; ALF76BECDSB9A827 CRC64;

Query Match
Best Local Similarity 24.8%; Score 148; DB 1; Length 528;
Matches 43; Conservative 23; Mismatches 37; Indels 26; Gaps 5;

QY 1 ARNPPIXWRLRQINSSESPPMYMLETSMGRSTOVQOSKDFIF-NTQIRLQISAG 60
DB 166 ABRFYISNRVYSIGTGFQPDAMISLENNWMLSRGVQSDPTFPMQ----- 216
QY 61 FTXSNVRD-----VSSLAIMFECGSRPFSSLDHPSPLLRSVVDANDVTC 109
DB 217 -VLTINIREPVIYDLSHPVAVLALMLFVCN--PPNA--NOSPILRSIYERK-ICS 270
QY 110 TXSEPTVRI 118
DB 271 SREPTVRI 279

RESULT 5
ABRC_ABRPR STANDARD; PRT: 562 AA.
AC P28590:
DB 01-DEC-1992 (Rel. 24, Created)
DB 15-DEC-1992 (Rel. 24, Last sequence update)
DB 15-JUN-2002 (Rel. 41, Last annotation update)
DE ABrin-c precursor [Contains: ABrin-c A chain (rRNA N-glycosidase)
DE (EC 3.2.2.22); ABrin-c B chain].
OS ABrus pectoratus (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidia I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=91266957; PubMed=2050149;
RA Wood K.A., Lord J.M., Mawrzynczak E.J., Platak M.;
RT "Preproabrin: genomic cloning, characterisation and the expression of
RT the A-chain in Escherichia coli.";
RL Eur. J. Biochem. 198;723-732(1991).
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RRNA. THE
CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE
CC BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entries requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL: X55667; CA939202.1;
DR PIR: S16022; S16022.
DR HSSP: P11140; 1ABR.
DR InterPro: IPR001574; RIP.
DR InterPro: IPR000772; Ricin_B_lectin.
DR Pfam: PF00161; RIP; 1.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR PRINTS: PRO0396; SHIGARICIN.
DR SMART: SMO0458; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIG_A_RICIN; 1.
KW Plant defense: Hydrolase; protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein;
KW Lectin; signal.
FT SIGNAL 1 34
FT CHAIN 35 285
FT PEPTIDE 286 295
FT CHAIN 296 562
FT DOMAIN 307 434
FT DOMAIN 437 561
FT REPEAT 317 359
FT REPEAT 360 400
FT REPEAT 403 435
FT REPEAT 448 483
FT REPEAT 487 526
FT REPEAT 529 562
FT ACT_SITE 198 198
FT DISULFID 281 303
FT DISULFID 320 339
FT DISULFID 363 380
FT DISULFID 451 464
FT DISULFID 490 507
FT MOD_RES 35 35
FT CARBOHYD 234 234
FT CARBOHYD 395 395
FT CARBOHYD 435 435
SQ SEQUENCE 562 AA; 62817 MW; 1FD0AB67D7BA6278 CRC64;

Query Match
Best Local Similarity 23.2%; Score 138; DB 1; Length 562;
Matches 43; Conservative 20; Mismatches 50; Indels 6; Gaps 4;

QY 1 ARNPPIXWRLRQINSSESPPMYMLETSMGRSTOVQOSKDFIF-NTQIRLQISAG 59
DB 200 ARFYISNRVYSIRGTARQPDPAAMLSLENNWMLSRGVQSDPTFPMNVLSSINQ 259
QY 60 NPYTXSNVRVYSSLAIMFECGSRPFSSLDHPSPLLRSVVDANDVCTYSEPTVRI 118
DB 260 PVAVDLSHPVAVLALMLFVCN--PPNA--NOSPILRSIYERK-ICSREPTVRI 313

RESULT 6
ABRB_ABRPR STANDARD; PRT: 527 AA.
AC Q06077; P81374;
DB 15-DEC-1998 (Rel. 37, Created)
DB 15-DEC-1998 (Rel. 37, Last sequence update)
DB 15-JUN-2002 (Rel. 41, Last annotation update)
DE ABrin-b precursor [Contains: ABrin-b A chain (rRNA N-glycosidase)
DE (EC 3.2.2.22); ABrin-b B chain].
OS ABrus pectoratus (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidia I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
OX NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=8421313;
RX Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;

```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U41299; AB39475.1; .
 DR PIR: S37382; S37382.
 DR PIR: S37383; S37383.
 DR InterPro: IPR001574; RIP.
 DR InterPro: IPR000772; Ricin_B_lectin.
 DR Pfam: PF00161; RIP; 2.
 DR Pfam: PF00652; Ricin_B_lectin; 6.
 DR PRINTS: PR00396; SHIGARICIN.
 DR SMART: SM00458; RICIN.
 DR PROSITE: PS00275; SHIGA_RICIN; 1.
 DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
 DR Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KM Glycoprotein; Lectin; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 297 NIGRIN B A CHAIN.
 FT CHAIN 298 563 NIGRIN B B CHAIN.
 FT DOMAIN 305 431 RICIN B-TYPE LECTIN 1.
 FT DOMAIN 434 559 RICIN B-TYPE LECTIN 2.
 FT REPEAT 316 356 1-ALPHA.
 FT REPEAT 357 397 1-BETA.
 FT REPEAT 400 432 1-GAMMA.
 FT REPEAT 445 482 2-ALPHA.
 FT REPEAT 486 524 2-BETA.
 FT REPEAT 527 554 2-GAMMA.
 FT ACT SITE 186 186 NIGRIN B-TYPE LECTIN 1.
 FT DISULFID 319 326 NIGRIN B-TYPE LECTIN 2.
 FT DISULFID 360 377 NIGRIN B-TYPE LECTIN 1.
 FT DISULFID 448 463 NIGRIN B-TYPE LECTIN 2.
 FT DISULFID 489 506 NIGRIN B-TYPE LECTIN 1.
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 39 39 K -> V (IN REF. 2).
 SQ SEQUENCE 563 AA; 62300 MW; F230CBE24621BF14 CRC64;

Query Match 18.1%; Score 108; DB 1; Length 563;
 Best Local Similarity 30.0%; Pred. No. 0.00012;
 Matches 36; Conservative 17; Mismatches 51; Indels 16; Gaps 5;

QY 1 ARNPFXWLRQINSRSGSPPMVMLEETSMGSRQSTQYQOSKDG1--FNTOIRLOIS 57
 DB 190 ARFPYIEQEVRSRLOQATSTFPNALMLSWENWSSMSLEIQAGNVSPFFGT----VQ 244
 QY 58 AGNEFVTVSNVRD-----VSSLAIMLFECGSRPSSLDIPSPLLLRVVAAND-VTQC 110
 DB 245 LLANDHTRLVDNFEELKKTITGIALILFRCSS-PSNDNAIRMLDLAGEDKRYKNDGRTY 303

RESULT 8
 RIPL_MOMCH STANDARD: PRT: 286 AA.
 AC P16094; P24697;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribosome-inactivating protein momordin I precursor (rRNA
 DE N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-MMC).
 OS Momordica charantia (Bitter melon) (Balsam pear).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Euphorbiales; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
 OC NCBI_TaxID=3673;

RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Seed;
 RX MEDLINE=91159486; PubMed=2001404;
 RA Ho W.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
 RT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
 RT protein.";
 RL Biochim. Biophys. Acta 1088:311-314(1991).
 RN (2)
 RP SEQUENCE OF 24-38.
 RC TISSUE-Seed;
 RX MEDLINE=89326691; PubMed=2753596;
 RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stierpe F., Sorla M.,
 RA Lappi D.;
 RT "N-terminal sequence of some ribosome-inactivating proteins.";
 RL Int. J. Pept. Protein Res. 33:263-267(1989).
 RN (3)
 RP SEQUENCE OF 24-70.
 RC TISSUE-Seed;
 RX MEDLINE=89005108; PubMed=3362509;
 RA Casellas P., Dussosoy D., Falasca A.I., Barbieri L.,
 RA Guillemot J.C., Ferrara P., Bolognesi A., Genini P., Stierpe F.;
 RT "Trichosanthin, a ribosome-inactivating protein from the seeds of
 RT trichosanthes kirilowii Maximowicz. Purification, partial
 RT characterization and use for preparation of immunotoxins.";
 RL Eur. J. Biochem. 176:581-588(1988).
 RN (4)
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=94356447; PubMed=8078985.
 RA Ren J., Wang Y., Dong Y., Stuart D.I.;
 RT "The N-glycosidase mechanism of ribosome-inactivating proteins
 RT implied by crystal structures of alpha-momorcharin.";
 RL Structure 2:7-16(1994).
 RN (5)
 RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
 RX MEDLINE=94192822; PubMed=8143869;
 RA Husain J., Tickie I.J., Wood S.P.;
 RT "Crystal structure of momordin, a type I ribosome inactivating
 RT protein from the seeds of Momordica charantia.";
 RL FEBS Lett. 342:154-158(1994).
 RN (6)
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=95344383; PubMed=7619070;
 RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
 RT "Studies on crystal structures, active-centre geometry and
 RT depurinating mechanism of two ribosome-inactivating proteins.";
 RL Biochem. J. 309:285-298(1995).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 CC -1- TYPE 1 RIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X57682; CA440869.1; .
 DR PIR: S14273; RLPUG.
 DR PIR: S16490; S16490.
 DR PDB: 1AHB; 22-JUN-94.
 DR PDB: 1AHB; 22-JUN-94.
 DR PDB: 1AHG; 22-JUN-94.
 DR PDB: 1MOM; 31-MAY-94.
 DR PDB: 1MRG; 07-FEB-95.
 DR PDB: 1MRH; 07-FEB-95.
 DR PDB: 1MRI; 07-FEB-95.
 DR GlycoSuiteDB: p16094; .
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP; 1.

```

DR PRINTS: P000396; SHIGARICIN.
DR PROSITE: P500275; SHIGA_RICIN: 1.
KW Glycoprotein; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 269 RIBOSOME-INACTIVATING PROTEIN MOMORDIN I.
FT PROPEP 270 286 MISSING IN MATURE PROTEIN.
FT ACT_SITE 183 183
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .).
/FTID-CAR.000082.
SQ SEQUENCE 286 AA; 31532 MW; E1B013ABEB216CF CRC64;
Query Match 15.4%; Score 92; DB 1; Length 286;
Best Local Similarity 34.1%; Pred. No. 0.0036;
Matches 28; Conservative 16; Mismatches 32; Indels 6; Gaps 3;
QY 1 ARRPPIKMLRQINSGSSPPNMVLEETSMGROSTOVQ--QSKDGIFNQRLOISA 58
DB 185 ARRYVIEOQIERARDE--VPSLATISLNSMSGLSKQIQLOAGNNGIFRPPIVLDNK 242
QY 59 GNEYXSNV--RDVYSLAIKL 78
DB 243 GNRVQITNVTSKVTSTNIDLL 264
RESULT 9
RIPA_PHYAM STANDARD; PRT; 294 AA.
AC 003464;
DR 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antiviral protein alpha precursor (PAP-alpha) (Ribosome-inactivating
protein) (rRNA N-glycosidase) (EC 3.2.2.22).
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Phytolacaceae; Phytolacca.
OX NCBI_TaxID=3527;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Seed, Leaf, and Root;
RX MEDLINE=9309240; PubMed=1281438;
RA Katoaka J., Habuka N., Masuta C., Miyano M., Koiwai A.;
RT "Isolation and analysis of a genomic clone encoding a pokeweed
antiviral protein."
RL Plant Mol. Biol. 20:879-886(1992).
RN (2)
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=95010127; PubMed=7925458;
RA Ago H., Katoaka J., Tsuge H., Hakuba N., Inagaki E., Noma M.,
MIYANO M.;
RT "X-ray structure of a pokeweed antiviral protein, coded by a new
genomic clone, at 0.23-nm resolution. A model structure provides a
suitable electrostatic field for substrate binding."
RL Eur. J. Biochem. 225:369-374(1994).
CC -1- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
SYNTHESIS IN VITRO. HAS ALSO BEEN SHOWN TO INHIBIT THE
REPLICATION OF MAMMALIAN VIRUSES. THE PROTEIN MAY PROVIDE A
MEANS OF CELLULAR SUICIDE UPON INVASION BY A VIRUS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Cell wall.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
TYPE 1 RIP SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb.ch/announce/)

```

```

CC or send an email to license@isb-sib.ch).
CC EMBL: D10600; BAA01451.1;
DR PIR: S28421; S28421.
DR PDB: IAPA; 3I-JAN-94.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP; 1.
DR PROSITE: P500275; SHIGA_RICIN: 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal; Cell wall; 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 285
FT PROPEP 286 294
FT ACT_SITE 199 199
FT DISULFID 58 282
FT DISULFID 108 130
FT STRAND 28 31
FT HELIX 32 34
FT HELIX 37 51
FT HELIX 57 58
FT STRAND 59 60
FT STRAND 61 63
FT TURN 67 68
FT STRAND 73 79
FT TURN 81 82
FT STRAND 85 91
FT TURN 92 94
FT TURN 97 104
FT STRAND 105 106
FT TURN 107 113
FT STRAND 114 115
FT TURN 119 129
FT HELIX 136 138
FT STRAND 141 141
FT STRAND 141 141
FT HELIX 147 154
FT TURN 155 155
FT HELIX 158 160
FT STRAND 163 163
FT HELIX 163 175
FT TURN 176 177
FT TURN 183 195
FT TURN 196 197
FT HELIX 198 202
FT HELIX 204 212
FT TURN 213 215
FT STRAND 218 218
FT HELIX 222 240
FT STRAND 242 242
FT TURN 243 244
FT STRAND 245 253
FT TURN 255 256
FT STRAND 259 264
FT HELIX 265 271
FT STRAND 275 275
SQ SEQUENCE 294 AA; 33069 MW; F2EC27724FA85596 CRC64;
Query Match 14.1%; Score 84; DB 1; Length 294;
Best Local Similarity 26.6%; Pred. No. 0.031;
Matches 21; Conservative 19; Mismatches 37; Indels 2; Gaps 1;
QY 1 ARRPPIKMLRQINSGSSPPNMVLEETSMGROSTOVQSKDGIFNQRLOISA 60
DB 201 ARRYVIEOQIERARDE--VPSLATISLNSMSGLSKQIQLOAGNNGIFRPPIVLDNK 258
QY 61 GNEYXSNV--RDVYSLAIKL 79
DB 259 KWIYLRVDIEPDVGLAKT 277
RESULT 10
RIP2_BRD1 STANDARD; PRT; 282 AA.
AC P98184; Q9S8J0.

```



```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein bryodin II precursor (rRNA N-
DE glycosidase) (EC 3.2.2.22) (BD2).
OS Bryonia dioica (Red bryony).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Bryonia.
NCBI_TaxID=3652;
RN [1]
RP SEQUENCE FROM N.A.
RA Siegel C.B., Gawlak S.L., Marguardt H.;
RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant
RL Bryonia dioica."
RL Patent number US597569, 28-JAN-1997.
RN [2]
RP SEQUENCE OF 22-42.
RC TISSUE-ROOT;
RX MEDLINE=95151812; PubMed=7849072;
RA Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
RA Marguardt H.;
RT "Characterization of ribosome-inactivating proteins isolated from
RT Bryonia dioica and their utility as carcinoma-reactive
RT immunoconjugates."
RT Blocom. Chem. 5:423-429(1994).
CC -I- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
CC PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: I34238; ; NOT_ANNOTATED_CDS.
DR HSSP: P09989; IMRJ.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN.1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW Multigene family; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 282 RIBOSOME-INACTIVATING PROTEIN BRYODIN II.
FT ACT_SITE 183 183 BY SIMILARITY.
FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;

Query Match 14.0%; Score 83.5; DB 1; Length 282;
Best Local Similarity 31.3%; Pred. No. 0.034; Indels 7; Gaps 4;
Matches 26; Conservative 15; Mismatches 35;

QY 1 ARFNPIXRRLROINSSESPNNMYMLETSMGROSTOVQ--QSKDGIFNTQIRLOI-- 56
DB 185 ARFYIIEORVSE--NWGTRKFPDPAPFLSLQNMAGSLSEQIQIAQTGTGGEFARPELVRTYS 242
QY 57 SAGNFVTXSNVRDYISSLAIMLF 79
DB 243 NTFPFVTNVN-SPVVKGIALLLY 264

RESULT 11
RIPS_PHYAM STANDARD; PRT; 261 AA.
AC P23339;
DT 01-NOV-1991 (Rel. 20, Created)

```

```

DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Antiviral protein S (PAP-S) (Ribosome-inactivating protein) (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Phytolacca americana (Common pokeweed) (Virginian pokeweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Phytolaccaceae; Phytolacca.
NCBI_TaxID=3527;
RN [1]
RP SEQUENCE.
RC TISSUE-Seed;
RX MEDLINE=91242096; PubMed=1368643;
RA Kung S.S., Kimura M., Funatsu G.;
RT "The complete amino acid sequence of antiviral protein from the seeds
RT of pokeweed (Phytolacca americana)."
RL Agric. Biol. Chem. 54:3301-3318(1990).
CC -I- FUNCTION: INHIBITS VIRAL INFECTION OF PLANTS, AND PROTEIN
CC SYNTHESIS IN VITRO.
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
DR PIR: JF0401; JF0401.
DR HSSP: O03464; IAPA.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN.1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin
FT ACT_SITE 175 175 BY SIMILARITY.
FT DISULFD 34 258
FT DISULFD 84 105
SQ SEQUENCE 261 AA; 29200 MW; D88B9962FE8399D CRC64;

Query Match 13.3%; Score 79; DB 1; Length 261;
Best Local Similarity 33.9%; Pred. No. 0.1;
Matches 20; Conservative 12; Mismatches 25; Indels 2; Gaps 1;

QY 1 ARFNPIXRRLROINSSESPNNMYMLETSMGROSTOVQSKDGIFNTQIRLOISAG 59
DB 177 ARFYIENQVATNFNRDPS--PNDKVLDEENMGKISTAIHNSKNGLAKPELEKNADG 233

RESULT 12
RIPS_TRIKI STANDARD; PRT; 289 AA.
AC P24478;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Ribosome-inactivating protein karasurin precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-Root tuber;
RX MEDLINE=97356562; PubMed=9212998;
RA Mizukami H., Iida K., Kondo T., Ogihara Y.;
RT "Cloning and bacterial expression of a gene encoding ribosome-
RT inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes
RT kirilowii var. japonica."
RL Biol. Pharm. Bull. 20:711-713(1997).
RN [2]
RP SEQUENCE OF 24-270.
RX MEDLINE=92005921; PubMed=1914000;
RA Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT "The complete amino acid sequence of an abortifacient protein,

```

```

RT   Karasurin."
RL   Chem. Pharm. Bull. 39:1244-1249(1991).
CC   -1- FUNCTION: ABORTION-INDUCING PROTEIN. IT INACTIVATES EUKARYOTIC
CC   60S RIBOSOMAL SUBUNITS.
CC   -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC   specific adenosine on the 28S rRNA.
CC   -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC   TYPE 1 RIP SUBFAMILY.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
CC   EMBL; AB000666; BAA21786.1; -.
CC   PIR; J00393; J00393.
CC   HSSP; P09989; 1MRJ.
CC   InterPro: IPR001574; RIP.
CC   Pfam: PF00161; RIP. 1.
CC   PRINTS: PR00396; SHIGARICIN.
CC   PROSITE: PS00275; SHIGA_RICIN. 1.
CC   Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
CC   Toxin; Signal.
CC   FT   SIGNAL. 1 21 POTENTIAL.
CC   FT   CHAIN 22 270 KARASURIN-C.
CC   FT   CHAIN 24 270 KARASURIN-A.
CC   FT   PROPEP 271 289 REMOVED IN MATURE FORM.
CC   ACT_SITE 183 183 BY SIMILARITY.
CC   FT   ACT_SITE 183 183
CC   FT   SEQUENCE 289 AA; 31704 MW; 883D3E242887B26 CRC64;
CC   SQ
Query Match 12.8%; Score 76.5; DB 1; Length 289;
Best Local Similarity 25.3%; Pred. No. 0.22;
Matches 25; Conservative 28; Mismatches 37; Indels 9; Gaps 4;

OY 1 ARENPIYRLRQINSSESPNMYMLEFTSGRSTQVO--QSKDGIFFNQIRLQISA 58
DB 185 ARYKFIFDQIGKRDV--KTLPSLAIITISLNSWSALSQIQIASTNNGQFETPVVLINQO 242
OY 59 GNEFTXSNVRD--VISSLAIMEFCSGRPSSSDHPSPL 95
DB 243 NORVTITNVDAGVYTSNIALIL---NRNNMAALDDDDVPM 278

RESULT 13
KATC_ARATH STANDARD: PRT; 754 AA.
ID KATC_ARATH STANDARD: PRT; 754 AA.
AC P46875; O9FH38;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Kinesin-like protein C.
GN KATC OR AT5G54670 OR K5F14.1 OR MRB17.18.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
OY 11
OY SEQUENCE FROM N.A.
OY STRAIN=cv Columbia;
OY MEDLINE=94355659; PubMed=8075402;
OY Mitsui H., Nakatani K., Yamaguchi-Shinozaki K., Shinozaki K.,
OY Nishikawa K., Takahashi H.;
OY "Sequencing and characterization of the kinesin-related genes katB
OY and katC of Arabidopsis thaliana.";
OY Plant Mol. Biol. 25:865-876(1994).
OY 12
OY SEQUENCE FROM N.A.
OY STRAIN=cv Columbia;
OY MEDLINE=20181125; PubMed=10718197;

```

```

RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
CC -1- FUNCTION: POSSIBLE ROLE IN MITOSIS.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS; A SMALL GLOBULAR
CC N-TERMINAL, A CENTRAL ALPHA-HELICAL COILED COIL AND A LARGE
CC GLOBULAR C-TERMINAL WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY
CC (IT HYDROLYZES ATP AND BINDS MICROTUBULES).
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. NCD
CC SUBFAMILY.
CC -1- CAUTION: Ref 2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D21138; BAA04674.1; -.
CC EMBL; AB022214; BAB09933.1; ALT_SEQ.
CC HSSP; P17119; 3KAR.
CC InterPro: IPR001752; kinesin_motor.
CC Pfam: PF00225; kinesin. 1.
CC PRINTS: PR00380; KINESINHEAVY.
CC SMART: SM00129; KISC. 1.
CC PROSITE: PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC PROSITE: PS50067; KINESIN_MOTOR_DOMAIN2; 1.
CC Motor protein; Microtubules; ATP-binding; Coiled coil;
CC Multigene family.
CC FT   DOMAIN 1 45 GLOBULAR.
CC FT   DOMAIN 46 393 COILED COIL.
CC FT   DOMAIN 398 727 KINESIN-MOTOR (BY SIMILARITY).
CC FT   NP_BIND 481 488 ATP (POTENTIAL).
CC FT   SEQUENCE 754 AA; 85030 MW; 76091CD5B5D9C531 CRC64;
CC SQ
Query Match 12.8%; Score 76; DB 1; Length 754;
Best Local Similarity 50.0%; Pred. No. 0.87;
Matches 15; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 28 ELFTSGRSTQVOYQSKDGIFFNQIRLQIS 57
DB 315 ELFTTSSQSTQIRQDLRLVNSRRLOYS 344

RESULT 14
RIPA_LUCFY STANDARD: PRT; 277 AA.
ID RIPA_LUCFY STANDARD: PRT; 277 AA.
AC O00465;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein Luffia luffin-alpha precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eucosids I; Cucurbitales; Cucurbitaceae; Luffa.
OX NCBI_TaxID=3670;
OY 11
OY SEQUENCE FROM N.A.
OY TISSUE=Seed;
OY MEDLINE=92286316; PubMed=1600156;
OY Katoaka J., Habuka N., Miyano M., Masuta C., Koizumi A.;
OY "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
OY inactivating protein from Luffa cylindrica.";
OY Plant Mol. Biol. 18:1199-1202(1992).
OY -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one

```

```

CC      specific adenosine on the 28S rRNA.
CC      -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC      TYPE 1 RIP SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb.slb.ch).
CC      -----
DR      EMBL: X62371; CAA44229.1; .
DR      PIR: S22494; S22494.
DR      HSSP: P16094; 1AHC.
DR      InterPro: IPR001574; RIP.
DR      Pfam: PF00161; RIP. 1.
DR      PRINTS: PR00396; SHIGARICIN.
DR      PROSITE: PS00275; SHIGA_RICIN; 1.
KW      Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
FT      Toxin; Signal.
FT      SIGNAL 1 19
FT      CHAIN 20 277
FT      ACT SITE 179 179
FT      SEQUENCE 277 AA; 30212 MW; EA17FC27998C25AC CRC64;
SQ
Query Match
Best Local Similarity 12.4%; Score 74; DB 1; Length 277;
Matches 21; Conservative 18; Mismatches 37; Indels 4; Gaps 2;
OY      1 AARNPXYMRLRQINSGESSPPNMYMLELTSWGRSTQVQ--OSKDGIPNTQIRLOISAGNFVTSNVDYISSLA1 76
DB      181 SRFKTEGILIRIS--KNQVPSLATISLENWMSLSKOILAQTNGTFTPVVIRDDK 238
OY      59 GNEVTSXSNVDYISSLA1ML 78
DB      239 GQRETVTVTSKVTKNIQL 258
RESULT 15
RIP1_CUCFI
ID      RIP1_CUCFI STANDARD; PRT; 286 AA.
AC      Q9FRX4;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)
DE      (EC 3.2.2.22).
OS      Cucumis figare1.
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX      NCBI_TaxID=131071;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Yamada T., Ohki S.T., Osaki T.;
RT      "Cloning and analysis of a cDNA coding a putative ribosome-
RT      inactivating protein from Cucumis figare1.";
RL      Plant Biotechnol. 17:337-340(2000).
CC      -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC      specific adenosine on the 28S rRNA.
CC      -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC      TYPE 1 RIP SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb.slb.ch).

```

```

DR      EMBL: AB045560; BAB19677.1; .
DR      HSSP: P16094; 1AHC.
DR      InterPro: IPR001574; RIP.
DR      Pfam: PF00161; RIP. 1.
DR      PRINTS: PR00396; SHIGARICIN.
DR      PROSITE: PS00275; SHIGA_RICIN; 1.
KW      Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
FT      Toxin; Signal.
FT      SIGNAL 1 21
FT      CHAIN 22 286
FT      ACT SITE 185 185
FT      CARBOHYD 103 103
FT      CARBOHYD 110 110
FT      CARBOHYD 252 252
FT      SEQUENCE 286 AA; 31771 MW; 4EFD4966504DAA1 CRC64;
SQ
Query Match
Best Local Similarity 12.4%; Score 74; DB 1; Length 286;
Matches 18; Conservative 13; Mismatches 29; Indels 2; Gaps 1;
OY      19 SSPPNMYMLELTSWGRSTQVQ--OSKDGIPNTQIRLOISAGNFVTSNVDYISSLA1 76
DB      203 SKVPDLAIAISLENWMSLSKOIAKSNNGQFQTVKIIINDKGLTEVTNVSLSVTKNI 262
OY      77 ML 78
DB      263 ML 264

```

Search completed: March 18, 2003, 08:25:43
Job time : 12.3819 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:23:18 ; Search time 15.8467 Seconds

(without alignments)
1547.297 Million cell updates/sec

Title: US-09-627-165E-14

Perfect score: 596

Sequence: 1 ARNPFXWRRLRQINGSSSS.....VVDANDVTCRXSEPTVRIV 119

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP yeast:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	408.5	68.5	565	10 Q8W243	Q8W243 viscum albu
2	387	64.9	531	10 Q8RXH6	Q8RXH6 viscum albu
3	324	54.4	249	10 Q8RXH7	Q8RXH7 viscum albu
4	193	32.4	592	10 Q8W2E7	Q8W2E7 Iris hollan
5	187	31.4	573	10 Q8W2E8	Q8W2E8 Iris hollan
6	179	30.0	549	10 Q9FV22	Q9FV22 cinamomum
7	179	30.0	580	10 Q94BW4	Q94BW4 cinamomum
8	179	30.0	580	10 Q94BW3	Q94BW3 cinamomum
9	179	30.0	581	10 Q94BW5	Q94BW5 cinamomum
10	178	29.9	541	10 Q41174	Q41174 ricinus com
11	157	26.3	382	10 Q8SA43	Q8SA43 abrus preca
12	139	23.3	528	10 Q06076	Q06076 abrus preca
13	128.5	21.6	603	10 Q9M653	Q9M653 polyonatum
14	128.5	21.6	604	10 Q9M654	Q9M654 polyonatum
15	122	20.5	547	10 Q9M659	Q9M659 abrus preca
16	121.5	20.4	570	10 Q41358	Q41358 sambucus n1

17	119.5	20.1	570	10 Q22415	Q22415 sambucus n1
18	109	18.3	563	10 Q04367	Q04367 sambucus n1
19	109	18.3	564	10 Q9AVR2	Q9AVR2 sambucus eb
20	107.5	18.0	251	10 Q96236	Q96236 abrus preca
21	107.5	18.0	252	10 Q38761	Q38761 abrus preca
22	107	17.3	563	10 Q94552	Q94552 sambucus n1
23	103	17.3	565	10 Q04071	Q04071 sambucus n1
24	103	17.3	566	10 Q04072	Q04072 sambucus n1
25	100.5	16.9	251	10 Q96235	Q96235 abrus preca
26	100.5	16.9	251	10 Q96237	Q96237 abrus preca
27	100.5	16.9	569	10 P93543	P93543 sambucus n1
28	98	16.4	300	10 Q04356	Q04356 Iris hollan
29	95	15.9	298	10 Q04358	Q04358 Iris hollan
30	93.5	15.7	252	10 Q38760	Q38760 abrus preca
31	91	15.3	300	10 Q04357	Q04357 Iris hollan
32	88	14.8	293	10 Q84552	Q84552 jatropa cu
33	85	14.3	293	10 Q8VYU0	Q8VYU0 jatropa cu
34	81	13.6	313	10 Q941G8	Q941G8 phytoacca
35	81	13.6	339	10 Q8RTA4	Q8RTA4 phytoacca
36	79	13.3	261	10 Q8S947	Q8S947 phytoacca
37	77.5	13.0	289	10 Q41216	Q41216 trichosanth
38	75	12.6	279	10 Q9AUE3	Q9AUE3 chenopodium
39	75	12.6	286	2 Q50196	Q50196 mycobacteri
40	75	12.6	302	10 P93077	P93077 clerodendru
41	74.5	12.5	556	4 Q9NV60	Q9NV60 homo sapien
42	74.5	12.5	1023	4 Q9ULF2	Q9ULF2 homo sapien
43	72.5	12.2	189	10 Q94KE4	Q94KE4 trichosanth
44	72	12.1	244	15 Q93GP4	Q93GP4 salmonele
45	72	12.1	237	10 Q9ATB3	Q9ATB3 phytoacca

ALIGNMENTS

RESULT 1

Q8W243 PRELIMINARY; PRT: 565 AA.
ID Q8W243;
AC Q8W243;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE VCA precursor.
OS viscum album subsp. coloratum
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_TaxID=159976;
RN [1]
RP SEQUENCE FROM N.A.
RA Park W.-B., Lyu S.;
RT "Cloning of Viscum album subsp. coloratum (Korean mistletoe).";
RL Biochem. Biophys. Res. Commun. 0:0-0(2002).
DR EMBL: AF369961; AAI0417.1;-.
DR InterPro: IPR001898; Gfam_pos_anchor.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; SHIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
KW Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 273 VCA ALPHA CHAIN.
FT CHAIN 309 565 VCA BETA CHAIN.
SQ SEQUENCE 565 AA: 62401 MW: 991E394DA005F11 CRC64:

Query Match 68.5%; Score 408.5; DB 10; Length 565;
Best Local Similarity 69.7%; Pred. No. 1,6e-38;
Matches 85; Conservative 10; Mismatches 24; Indels 3; Gaps 1;

OVERALL	3	PRELIMINARY;	PRT;	249	AA.
OBXHX7	OBXHX7				
ID	OBXHX7				
AC	OBXHX7				
DT	01-JUN-2002 (TREMBLrel_21, Created)				
DT	01-JUN-2002 (TREMBLrel_21, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel_21, Last annotation update)				
DE	lectin chain A isoform 2 (Fragment).				
OS	Viscum album (European mistletoe).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
OC	Santalales; Viscaceae; Viscum.				
OX	NCBI_TaxID=3972;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=LEAF;				
RA	paramasium M., Srihivasan A., Singh T.P.;				
RT	"Viscum Album (Indian) mRNA for Mistletoe lectin Chain A, isoform 2."				
RL	Submitted (Feb-2002) to the EMBL/Genbank/DBJ databases.				
DR	EMBL: A1081148; A1487005.1; -				
FT	NON_TER	1	1	249	249

RESULT	5	
Q8WZ88		
Q8WZ88	PRELIMINARY;	PRT; 573 AA.
ID	Q8WZ88;	
DT	01-MAR-2002 (TRMBLrel. 20, Created)	
DT	01-MAR-2002 (TRMBLrel. 20, Last sequence update)	
DT	01-JUN-2002 (TRMBLrel. 21, Last annotation update)	
DE	Ribosome-inactivating protein Irar (Fragment).	
GN	LECIRAR.	
GS	Iris hollandica (Dutch iris).	

```

OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Iridaceae;
OX NCBI_TaxID=35876;
RN [1]
RP SEQUENCE FROM N.A.
RA Van Damme E.J.M., Peumans W.J.;
RT "Iris (Iris hollandica var. Professor Blaauw) plants express both type
RT 1 and type 2 ribosome-inactivating proteins in bulb tissue.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF256084; AL55093.1; -
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PRO0396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 573 AA; 63759 MW; 1414A3B9AECDAF5C CRC64;

Query Match 31.4%; Score 187; DB 10; Length 573;
Best Local Similarity 37.7%; Pred. No. 5.6e-13;
Matches 52; Conservative 19; Mismatches 43; Indels 24; Gaps 6;

QY 1 ARFNPXWRLRR--QINSGSSPPNMYMLETSMGROSTOVQOOSKDGFNTQIRLOIS 57
DB 191 ARFNPXWRLRRNIIQVGDVRSFPGAGMDLETNMEPLSERVQESNCGVFNRLTQTS 250
QY 58 AGNFTV---XSNRVYISSLAIMLFECSSGRPFSSLD-----HPSPLL-----LRSV 100
DB 251 --NETIDIVNAQARQVCGALLLFACKAR--QSLQALRPHDSPLVPLITLDLVNVRSM 306
QY 101 VDAANDVTCIXSEPTVRI 118
DB 307 LDIVEDDTCPLSEPTMRI 324

RESULT 6
ID Q9FV22 PRELIMINARY; PRT; 549 AA.
AC Q9FV22;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE rRNA-glycosidase (EC 3.2.2.22) (Fragment).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Xie L., Liu W.-Y., Wang E.-D.;
RT "Molecular cloning of cinnamomin A", B-chain and the expression,
RT purification, characterization and mutagenesis of the A-chain.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOLYDOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL; AF229548; AF68978.2; -
DR HSSP: P02879; 2AA1.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 5.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PRO0396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR Hydroxylase; Toxin.
FT NON_TER 1
SQ SEQUENCE 549 AA; 60648 MW; 02607FE607CA44B0 CRC64;

```

```

Query Match 30.0%; Score 179; DB 10; Length 549;
Best Local Similarity 38.9%; Pred. No. 4.5e-12;
Matches 51; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 RNPPIXRRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDGIFNTQIRLOISAGN 60
DB 170 RNPPIXRRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDGIFNTQIRLOISAGN 229
QY 61 FVYXSNVRD-VISSLAIMLFEC--SGRPSS--LDHPSPLLRSVY-----DAANDV 107
DB 230 FVYXSNVRD-VISSLAIMLFECSTDRASSDQFIDH--MLMIRPLVDAVATDANDND 287
QY 108 TCTXSEPTVRI 118
DB 288 TCADPEPTVRI 298

RESULT 7
ID Q94BW4 PRELIMINARY; PRT; 580 AA.
AC Q94BW4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE rRNA-glycosidase precursor (EC 3.2.2.22).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOLYDOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL; AF039802; AK82459.1; -
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
DR Hydroxylase; Signal; Toxin.
FT SIGNAL 1
FT CHAIN 33 580
FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN II.
SQ SEQUENCE 580 AA; 64265 MW; 37E4289ECC0CBFF CRC64;

Query Match 30.0%; Score 179; DB 10; Length 580;
Best Local Similarity 40.5%; Pred. No. 4.8e-12;
Matches 53; Conservative 16; Mismatches 42; Indels 16; Gaps 6;

QY 2 RNPPIXRRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDGIFNTQIRLOISAGN 60
DB 202 RNPPIXRRLRQINSGSSPPNMYMLETSMGROSTOVQOOSKDGIFNTQIRLOISAGN 261
QY 61 FVYXSNVRD-VISSLAIMLFEC--SGRPSS--LDHPSPLLRSVY-----DAANDV 107
DB 262 FVYXSNVRD-VISSLAIMLFECSTDRASSDQFIDH--MLMIRPLVDAVATDANDND 319
QY 108 TCTXSEPTVRI 118
DB 320 TCADPEPTVRI 330

RESULT 8
ID Q94BW3 PRELIMINARY; PRT; 580 AA.
AC Q94BW3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```

```

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE rRNA -glycosidase precursor (EC 3.2.2.22).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxId=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu M.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomn proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: AY039803; AAK82460.1; -.
DR InterPro: IPR000772; RIBCIN_B_lectin.
DR Pfam: PF00652; RIBCIN_B_lectin; 6.
DR PROSITE: PS50231; RIBCIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 580 TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOIN III.
SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7EB558 CRC64;

Query Match 30.0%; Score 179; DB 10; Length 580;
Best Local Similarity 38.9%; Pred. No. 4.8e-12;
Matches 51; Conservative 17; Mismatches 47; Indels 16; Gaps 5;

QY 2 RNPPIKWLRLRQINSSESPPMYMLEETSGRSTOVQOOS-KDGIPTQIRLQISAGN 60
DB 202 RFRLEFYRVRSGISRAEFMRPDPALSLKMKWALSNAVOOSNGGVFSSPVELRSISNK 261
QY 61 FVTXSNVRD-VISSLAIMFECSGRPFSS---LDHPSLLRLRSV-----DAANDY 107
DB 262 PTVGVSVDRIYSLAIFLCRSTDRASSDQFIDH--LIMIRPLDVAVATDADND 319
QY 108 TCTXSEPTVRI 118
DB 320 TCADEPPTVRI 330

RESULT 9
Q94BW5 PRELIMINARY; PRT; 581 AA.
AC Q94BW5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE rRNA -glycosidase precursor (EC 3.2.2.22).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxId=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu M.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomn proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: AY039801; AAK82458.1; -.
DR InterPro: IPR000772; RIBCIN_B_lectin.
DR Pfam: PF00652; RIBCIN_B_lectin; 5.
DR Pfam: PF00161; RIP; 1.

```

```

DR PROSITE: PS50231; RIBCIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 581 TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOIN I.
SQ SEQUENCE 581 AA; 64215 MW; 6EBF5FB8FBA3D196 CRC64;

Query Match 30.0%; Score 179; DB 10; Length 581;
Best Local Similarity 38.9%; Pred. No. 4.8e-12;
Matches 51; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 RNPPIKWLRLRQINSSESPPMYMLEETSGRSTOVQOOS-KDGIPTQIRLQISAGN 60
DB 202 RFRLEFYRVRSGISRAEFMRPDPALSLKMKWALSNAVOOSNGGVFSSPVELRSISNK 261
QY 61 FVTXSNVRD-VISSLAIMFECSGRPFSS---LDHPSLLRLRSV-----DAANDY 107
DB 262 PTVGVSVDRIYSLAIFLCRSTDRASSDQFIDH--LIMIRPLDVAVATDADND 319
QY 108 TCTXSEPTVRI 118
DB 320 TCADEPPTVRI 330

RESULT 10
Q41174 PRELIMINARY; PRT; 541 AA.
AC Q41174;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
OS Ricinus communis (Castor bean).
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eumids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxId=3986;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92338377; PubMed=1633311;
RA Roberts L.M., Tregear J.W., Lord J.M.;
RT "Molecular cloning of ricin.";
RL Targeted Diagn. Ther. 7:81-97(1992).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: S40366; AAB22582.1; -.
DR HSSP: P02879; 1BR6.
DR InterPro: IPR000772; RIBCIN_B_lectin.
DR InterPro: IPR001574; RIP.
DR InterPro: IPR001400; Somatotropin.
DR Pfam: PF00652; RIBCIN_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RIBCIN_2.
DR PROSITE: PS50231; RIBCIN_B_LECTIN; 2.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; UNKNOWN_1.
KW Hydrolase; Toxin.
FT NON TER 1 1
SQ SEQUENCE 541 AA; 60281 MW; 2B7B2CDEF1FE29D9 CRC64;

Query Match 29.9%; Score 178; DB 10; Length 541;
Best Local Similarity 37.0%; Pred. No. 5.7e-12;
Matches 44; Conservative 19; Mismatches 52; Indels 4; Gaps 1;

QY 1 ARNPPIKWLRLRQINSSESPPMYMLEETSGRSTOVQOOSKDGIPTQIRLQISAGN 60
DB 179 ARFOYLEGEMRTIRIRNRRSAPDPVITLLENSWGRSLTAIOESNGCAFASPIOLGRNRS 238
QY 61 FVTXSNVRDVISSLAIMFECSGRPFSSLDHPSELLRSVVAANDVYCTXSEPTVRI 119
DB 239 KFSYVDVSLIPLIALMKVRCAPPSOPS---LIRVYVNFNADVCMDEPPTVRI 293

```


RESULT 11

Q8SA43 PRELIMINARY; PRT; 382 AA.

AC 08SA43; TREMBLrel. 21, Created

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Abirin isoform G (Fragment).

OS *Abrus precatorius* (Indian licorice) (Crab's eye).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.

OX NCBI_TaxID=3816;

RP SEQUENCE FROM N.A.

RC TISSUE=LEAF;

RA Cook J.P., Roberts L.M., Lord M.;

RT "New Isoform of Abirin - Abirin G";

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF479626; AAL7434.1; -

FT CHAIN 1 >105 ABRIN A CHAIN.

FT NON_TER 1 382 ABRIN B CHAIN.

FT CHAIN 116 >382

FT NON_TER 382

SO SEQUENCE 382 AA; 42743 MW; B08AB34181ADZEE CRC64;

Query Match 26.3%; Score 157; DB 10; Length 382;

Best Local Similarity 36.2%; Pred. No. 1e-09;

Matches 47; Conservative 20; Mismatches 35; Indels 28; Gaps 6;

QY 1 ARFNPIYKRLRQINSGESSPPNMYMLEETSMGRQSTQVOQSDGIFNPIRLQISAGN 60

DB 20 ARFRYSNRVGVSIKGTAFQDPDPMLSLENWMDLSGGVQVQADQDF-----PN 69

QY 61 FVTXSNVND-----VSSLAIMLFECSGRPFSSLDHPSPLLRSVVDANDVTC 109

DB 70 DVTLRNTRDEPVYVNSLHPYVAVALMLFVCN--PPNA--NOSPILRSIVEKSD--TC 123

QY 110 TXS-EPTVRI 118

DB 124 VPPIEPTVRI 133

RESULT 12

Q06076 PRELIMINARY; PRT; 528 AA.

AC Q06076; TREMBLrel. 01, Created

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).

OS *Abrus precatorius* (Indian licorice) (Crab's eye).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.

OX NCBI_TaxID=3816;

RP SEQUENCE FROM N.A.

RM MEDLINE=93132798; PubMed=8421313;

RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;

RT "Primary structure of three distinct isobabins determined by cDNA

RT sequencing: conservation and significance.";

RL J. Mol. Biol. 229:263-267(1993)

CC -1- CATALYTIC ACTIVITY: ENDORHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE

CC SPECIFIC ADENOSINE ON THE 28S RNA.

CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.

DR EMBL; M98346; AAA32626.1; -

DR HSSP; P11140; IABR.

DR InterPro; IPR000772; R1Cln_B_lectin.

DR InterPro; IPR001574; R1P.

DR Pfam; PF00652; R1Cln_B_lectin; 6.

DR Pfam; PF00161; R1P; 1.

DR PRINTS; PR00396; SHIGARICIN.

DR SMART; SM00458; R1CIN; 2.

DR PROSITE; PS00231; R1CIN_B_LECTIN; 2.

DR PROSITE; PS00275; SHIGA_R1CIN; UNKNOWN_1.

FT NON_TER 1 528

FT NON_TER 528

SO SEQUENCE 528 AA; 58870 MW; 62ED42FB8FFE60F8 CRC64;

Query Match 23.3%; Score 139; DB 10; Length 528;

Best Local Similarity 36.1%; Pred. No. 1.7e-07;

Matches 43; Conservative 21; Mismatches 49; Indels 6; Gaps 4;

QY 1 ARFNPIYKRLRQINSGESSPPNMYMLEETSMGRQSTQVOQSDGIFNPIRLQISAG 59

DB 166 ARFRISNRVGVSIKGTAFQDPDPMLSLENWMDLSGGVQVQADQDFPNVILSSINRQ 225

QY 60 NEVTSXNVNDVSSLAIMLFECSGRPFSSLDHPSPLLRSVVDANDVTCXSEPTVRI 118

DB 226 PVVVDLSHPYVAVALMLFVCN--PPNA--NOSPILRSIVEKSD--ICSRYEPTVRI 279

RESULT 13

Q09M53 PRELIMINARY; PRT; 603 AA.

AC Q09M53; TREMBLrel. 15, Created

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE rRNA -glycosidase (EC 3.2.2.22).

GN R1P.

OS Polygonatum multiflorum.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;

OC Convallariaceae; Polygonatum.

OX NCBI_TaxID=45371;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20247177; PubMed=10785398;

RA Van Damme E.J., Hao Q., Charels D., Barre A., Rouge P., Van Leuven F.,

RA Peunmans W.D.;

RT "Characterization and molecular cloning of two different type 2

RT ribosome-inactivating proteins from the monocotyledonous plant

RT Polygonatum multiflorum.";

RL Eur. J. Biochem. 267:2746-2759(2000).

CC -1- CATALYTIC ACTIVITY: ENDORHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE

CC SPECIFIC ADENOSINE ON THE 28S RNA.

CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.

DR EMBL; AF213984; AAF37219.1; -

DR HSSP; P02879; 2AAL.

DR InterPro; IPR000772; R1Cln_B_lectin.

DR InterPro; IPR001574; R1P.

DR Pfam; PF00652; R1Cln_B_lectin; 6.

DR Pfam; PF00161; R1P; 1.

DR PRINTS; PR00396; SHIGARICIN.

DR SMART; SM00458; R1CIN; 2.

DR PROSITE; PS00231; R1CIN_B_LECTIN; 2.

DR PROSITE; PS00275; SHIGA_R1CIN; UNKNOWN_1.

DR PROSITE; PS00275; SHIGA_R1CIN; UNKNOWN_1.

KW Hydroxylase; Toxin.

SO SEQUENCE 603 AA; 66733 MW; 88B61C17CA431BIE CRC64;

Query Match 21.6%; Score 128.5; DB 10; Length 603;

Best Local Similarity 31.4%; Pred. No. 3.3e-06;

Matches 44; Conservative 20; Mismatches 49; Indels 27; Gaps 7;

QY 1 ARFNPIYKRLRQINSGESSPPNMYMLEETSMGRQSTQVOQSDGIFNPIRLQISAG 55

DB 210 ARFRHIEERVRSRFTADHQLFRPDGLMLSMENWPSMSLVSIGSITGVQVQ 269

QY 56 -----ISAGNFVYXSNVNDVSSLAIMLFECSGRPFSSLDHPSPLLRSVVDAA 104

DB 270 DESFOPLRVNDNFNLSR-----YTNVALLLFR-CHPRATAGTSSTTPMAAQILRMVVDYL 324

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:00:53 ; Search time 49,0704 Seconds

(Without alignments)
757.625 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1425

Sequence: 1 HQTGDEYFRFTLRDYS.....NSGAVDVTCTASEPTVRIV 279

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1389	97.5	551	23	ABR79450
2	1269.5	89.1	531	20	AAV25979
3	1269.5	89.1	532	20	AAV25962
4	1269.5	89.1	564	18	AAW10021
5	1269.5	89.1	564	20	AAW90127
6	1211	85.0	533	20	AAV25970
7	1211	85.0	533	20	AAV25973
8	1211	85.0	533	20	AAV25976
9	1155	81.1	254	22	AAW47090
10	1152	80.8	256	20	AAV25981

11	1152	80.8	256	20	AAV25984
12	1131	79.4	252	19	AAW64661
13	1131	79.4	253	18	AAW10022
14	1131	79.4	253	20	AAW90125
15	1127	79.1	252	19	AAW64659
16	1119	78.5	254	20	AAV25980
17	1119	78.5	254	20	AAV25983
18	1090	76.5	254	22	AAW47091
19	1080.5	75.8	255	20	AAV25971
20	1080.5	75.8	255	20	AAV25974
21	1080.5	75.8	255	20	AAV25977
22	1067	74.9	256	22	AAW47092
23	885	62.1	174	22	AAW47097
24	474	33.3	251	21	AAW19267
25	466.5	32.7	250	18	AAW25137
26	466.5	32.7	250	18	AAW21700
27	461	32.4	252	16	AAW81976
28	458	32.1	252	16	AAW81977
29	452	31.7	253	16	AAW81978
30	425.5	29.9	332	8	AAW70097
31	425.5	29.9	332	8	AAW70098
32	425.5	29.9	332	10	AAW93639
33	425.5	29.9	332	11	AAW06554
34	425.5	29.9	562	10	AAW90079
35	425.5	29.9	576	8	AAW70326
36	425.5	29.9	576	18	AAW25787
37	425.5	29.9	576	20	AAV55892
38	425.5	29.9	576	21	AAV78592
39	425.5	29.9	576	22	AAW78301
40	425.5	29.9	576	22	AAW78302
41	423.5	29.7	574	10	AAW94793
42	423	29.7	540	18	AAW25143
43	423	29.7	540	18	AAW21706
44	422.5	29.6	565	6	AAW50166
45	422.5	29.6	565	7	AAW60240

ALIGNMENTS

RESULT 1	
ABR79450	
ID	ABR79450 standard; Protein; 551 AA.
AC	ABR79450:
XX	
DT	08-JUL-2002 (first entry)
XX	
DE	Galactose-recognising mistletoe lectin.
XX	
KW	Mistletoe; galactose-recognising mistletoe lectin; MLIII.
XX	
OS	Viscum album.
XX	
FH	Key
FT	Misc-difference 223
FT	/note- "Encoded by ATG"
FT	Misc-difference 251
FT	/note- "Encoded by TTP"
FT	Misc-difference 344
FT	/note- "Encoded by TCC"
FT	Misc-difference 380
FT	/note- "Encoded by GCC"
FT	Misc-difference 448
FT	/note- "Encoded by GTG"
PN	DE10044027-A1.
XX	
PD	14-MAR-2002.
XX	
PF	06-SEP-2000; 2000DE-1044027.
XX	
PR	06-SEP-2000; 2000DE-1044027.

Mistletoe lectin A
Mistletoe rMLA var
Prepro mistletoe I
Mistletoe ML A-cha
Mistletoe rMLA pro
Mistletoe lectin A
Mistletoe lectin A
A-chain isoform fo
Mistletoe lectin A
Mistletoe lectin A
Mistletoe lectin A
A-chain isoform fo
Amino acid sequenc
Arlin A A-chain ri
Arlin A A-chain.
Recombinant arlin
Recombinant arlin
Riclin A. Escheric
Sequence of Riclinu
Riclin A encoded by
Riclin A gene produ
Riclin D. Riclinus
Sequence of Riclinu
Castorbean riclin.
Castor bean riclin
Riclinus communis r
Castor bean prepro
DNA sequence of ri
Castor oil plant A
R. communis agglut
Sequence of prepro
PreproRiclin. Ricl

```
XX (VISC-) VISCUM AG.
PA
XX
XX Kieff S;
XX
XX WPI; 2002-316737/36.
XX N-PSDB; ABL56947.
XX
PT New nucleic acid encoding preprotein of mistletoe lectin, useful as
PT diagnostic and therapeutic agents, also encodes polypeptide
XX
XX Claim 1; Fig 1; 6pp; German.
XX
CC The invention relates to a nucleic acid molecule (ABL56947) that encodes
CC a preprotein (AB879450) which, after maturation, has the biological
CC activity of the galactose-recognising mistletoe lectin (MLII). The MLII
CC encoding nucleic acid molecule, primers specific to it or complements of
CC it, and encoded (oligomeric) polypeptides are useful as diagnostic and
CC therapeutic agents.
XX
SQ Sequence 551 AA;
Query Match 97.5%; Score 1389; DB 23; Length 551;
Best Local Similarity 97.5%; Pred. No. 1,6e-133;
Matches 272; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 1 HQTGDEYFRFTLLRDYSSGFSFNEIPILROSTIPVSDAQRVYVELTNOGDSITPA 60
DB 43 HQTGGEYFRFTLLRDYSSGFSFNEIPILROSTIPVSDAQRVYVELTNOGDSITPA 102
OY 61 IDVTNLVYVAYQAGDQSYFLRDAPDGAERHLEFTGTTRSSLPTGSGYTLDERYAGHRDQIP 120
DB 103 IDVTNLVYVAYQAGDQSYFLRDAPDGAERHLEFTGTTRSSLPTGSGYTLDERYAGHRDQIP 162
OY 121 LGIEELIOSVSAALRYPGSGTRQAARSLITLIOMISEAARFNPFWRARQYINSGESFLPD 180
DB 163 LGIEELIOSVSAALRYPGSGTRQAARSLITLIOMISEAARFNPFWRARQYINSGESFLPD 222
OY 181 MYMLELETSWGOQSTOVQOQSTGDFVNNPFRIGISTGNEFVTLNVRDVIASLAIMLFVCRD 240
DB 223 VYMLELETSWGOQSTOVQOQSTGDFVNNPFRILAISTGNEFVTLNVRDVIASLAIMLFVCRD 282
OY 241 RPSSSDVRYWPLVIRPVLENSGAVDDYCTASEPTVRIY 279
DB 283 RPSSSDVRYWPLVIRPVLENSGAVDDYCTASEPTVRIY 321
RESULT 2
AAV25979
ID AAV25979 standard; Protein; 531 AA.
XX
XX AAV25979;
XX
DT 18-OCT-1999 (first entry)
XX
DE Mistletoe lectin I protein fragment.
XX
XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
XX ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX cancer; cytotoxicity; antigen; isoform; lectin I.
XX
OS Viscum album.
XX
XX DEL19804210-A1.
XX
XX 12-AUG-1999.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
PA
```

```
XX
XX Morris P, Stiefel T, Voelter W, Welters P;
XX
XX WPI; 1999-445335/38.
XX N-PSDB; AAZ09103.
XX
PT Preparation of mistletoe lectins in heterologous systems,
PT particularly for use as anticancer agents and immunostimulants
XX
XX Claim 7; Fig 1B; 78pp; German.
XX
CC This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MIA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
XX represents a mistletoe lectin I protein fragment.
SQ Sequence 531 AA;
Query Match 89.1%; Score 1269.5; DB 20; Length 531;
Best Local Similarity 88.9%; Pred. No. 2.8e-121;
Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;
OY 1 HQTGDEYFRFTLLRDYSSGFSFNEIPILROSTIPVSDAQRVYVELTNOGDSITPA 60
DB 10 HQTGGEYFRFTLLRDYSSGFSFNEIPILROSTIPVSDAQRVYVELTNOGDSITPA 69
OY 61 IDVTNLVYVAYQAGDQSYFLRDAPDGAERHLEFTGTTRSSLPTGSGYTLDERYAGHRDQIP 120
DB 70 IDVTNLVYVAYQAGDQSYFLRDAPDGAERHLEFTGTTRSSLPTGSGYTLDERYAGHRDQIP 129
OY 121 LGIEELIOSVSAALRYPGSGTRQAARSLITLIOMISEAARFNPFWRARQYINSGESFLPD 180
DB 130 LGIDQLIOSVYALRFPGSGTRQAARSLITLIOMISEAARFNPFWRARQYINSGESFLPD 189
OY 181 MYMLELETSWGOQSTOVQOQSTGDFVNNPFRIGISTGNEFVTLNVRDVIASLAIMLFVCRD 240
DB 190 VYMLELETSWGOQSTOVQOQSTGDFVNNPFRILAIPPGNEFVTLNVRDVIASLAIMLFVCRD 249
OY 241 RPSSSDVRYWPLVIRPVLENSGAVDDYCTASEPTVRIY 279
DB 250 RPSSSDVRYWPLVIRPVI-----ADVTGSASEPTVRIY 283
RESULT 3
AAV25982
ID AAV25982 standard; Protein; 532 AA.
XX
XX AAV25982;
XX
DT 18-OCT-1999 (first entry)
XX
DE Mistletoe lectin I (variant) protein fragment.
XX
XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
XX ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX cancer; cytotoxicity; antigen; isoform; lectin I.
XX
OS Viscum album.
XX
XX DEL19804210-A1.
XX
XX 12-AUG-1999.
XX
XX
```

PF 03-FEB-1998: 98DE-1004210.
 XX 03-FEB-1998: 98DE-1004210.
 PR (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PA
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI, 1999-445335/38.
 DR N-PSDB; AAO09106.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 particularly for use as anticancer agents and immunostimulants
 XX
 PS Disclosure; Fig 4B; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of a mistletoe lectin I protein variant.
 XX
 SQ Sequence 532 AA;
 Query Match 89.1%; Score 1269.5; DB 20; Length 532;
 Best Local Similarity 88.9%; Pred. No. 2.8e-121;
 Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;
 QY 1 HQTGDEYFRFTLLRDYVSSGFSNEIPLLRQSTIPVSDAQRFLVLELTNOGDSITAA 60
 DB 10 HQTGDEYFRFTLLRDYVSSGFSNEIPLLRQSTIPVSDAQRFLVLELTNOGDSITAA 69
 QY 61 IDVTNLVYVAYOAGDOSYFLRDAPDGAERHLFTGTRSSLPFTGSYTDLERVAGHRDQIP 120
 DB 70 IDVTNLVYVAYOAGDOSYFLRDAPDGAERHLFTGTRSSLPFTGSYTDLERVAGHRDQIP 129
 QY 121 LGIEELIQSVSLARYPGGSTRQAARSLILIQMISEAARFNPIFWARQYINSGESFLPD 180
 DB 130 LGIDQLIOSYVALRFPGGSTRQAARSLILIQMISEAARFNPIFWARQYINSGASFLPD 189
 QY 181 MYMLEETSMGQOSTOVQOSTDGVFNPNFRIGISTGNFVTLISNVBDVYASLAIMLFCRD 240
 DB 190 VYMLEETSMGQOSTOVQOSTDGVFNPNFRILALIPGNEFVTLINVDVYASLAIMLFCVGE 249
 QY 241 RPSSSDVRYWPLYRPLENSGAVDDYCTASEPTVRIY 279
 DB 250 RPSSSDVRYWPLYRPIV-----ADVTCASASEPTVRIY 283
 RESULT 4
 AAM10021
 ID AAM10021 standard; Protein; 564 AA.
 XX
 AC AAM10021;
 XX
 DT 18-DEC-1997 (first entry)
 XX
 DE Prepro mistletoe lectin.
 XX
 KW Mistletoe; lectin; cytotoxic; A chain; B chain; dimer.
 XX
 OS Viscum album.
 XX
 PN EP751221-A1.
 XX

PD 02-JAN-1997.
 XX
 PF 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 95EP-0109949.
 XX
 PA (MADU) MADAU KOELN AG.
 XX
 PI Baur A, Eck J, Lentzen H, Zinke H;
 XX WPI, 1997-054678/06.
 DR N-PSDB; AAT70473.
 XX
 PT Nucleic acid encoding pre-pro form of mistletoe lectin - for
 PT therapeutic or diagnostic use
 XX
 PS Claim 12; Fig 4C; 30pp; German.
 XX
 CC Mistletoe lectin is a cytotoxic agent that has been used for tumour
 CC therapy. It can be used in immunotoxins and medicaments. Nucleic
 CC acid fragments can be used in diagnostic methods. Mistletoe lectin (
 CC AAT70473) comprises an A chain (AAT70474) and a B chain (AAT70475).
 XX
 SQ Sequence 564 AA;
 Query Match 89.1%; Score 1269.5; DB 18; Length 564;
 Best Local Similarity 88.9%; Pred. No. 3e-121;
 Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;
 QY 1 HQTGDEYFRFTLLRDYVSSGFSNEIPLLRQSTIPVSDAQRFLVLELTNOGDSITAA 60
 DB 43 HQTGDEYFRFTLLRDYVSSGFSNEIPLLRQSTIPVSDAQRFLVLELTNOGDSITAA 102
 QY 61 IDVTNLVYVAYOAGDOSYFLRDAPDGAERHLFTGTRSSLPFTGSYTDLERVAGHRDQIP 120
 DB 103 IDVTNLVYVAYOAGDOSYFLRDAPDGAERHLFTGTRSSLPFTGSYTDLERVAGHRDQIP 162
 QY 121 LGIEELIQSVSLARYPGGSTRQAARSLILIQMISEAARFNPIFWARQYINSGESFLPD 180
 DB 163 LGIDQLIOSYVALRFPGGSTRQAARSLILIQMISEAARFNPIFWARQYINSGASFLPD 222
 QY 181 MYMLEETSMGQOSTOVQOSTDGVFNPNFRIGISTGNFVTLISNVBDVYASLAIMLFCRD 240
 DB 223 VYMLEETSMGQOSTOVQOSTDGVFNPNFRILALIPGNEFVTLINVDVYASLAIMLFCVGE 282
 QY 241 RPSSSDVRYWPLYRPLENSGAVDDYCTASEPTVRIY 279
 DB 283 RPSSSEVRWPLYRPIV-----ADVTCASASEPTVRIY 316
 RESULT 5
 AAM90127
 ID AAM90127 standard; Protein; 564 AA.
 XX
 AC AAM90127;
 XX
 DT 30-APR-1999 (first entry)
 XX
 DE Mistletoe lectin prepro-protein.
 XX
 KW ML; mistletoe; lectin; ML; transgenic plant; glycosylation;
 KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;
 KW cancer.
 XX
 OS Viscum album.
 XX
 PN EP84388-A1.
 XX
 PD 16-DEC-1998.
 XX
 PF 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 95EP-0109949.
 XX

PR 26-JUN-1995; 98EP-0105660.
XX
XX (MADU) MADAU KOELN AG.
XX
PI Baur A, Eck J, Lentzen H, Zinke H;
XX WPI: 1999-026582/03.
DR N-PSDB; AAV74182.
XX
PT New transgenic plant expressing mistletoe lectin - useful for
PT producing recombinant lectin in e.g. cancer diagnosis and therapy
XX
PS Claim 1a; Fig 4c; 30pp; German.
XX
XX This invention describes a novel transgenic plant transformed with a
CC vector capable of encoding a mistletoe (Viscum album) lectin
CC preproprotein or a biologically active fragment. The specification
CC also describes a polypeptide produced by a plant where the polypeptide
CC exhibits at least one enzymatic modification other than the glycosylation
CC that occurs in Viscum album or the polypeptide is a fusion protein, a
CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the
CC polypeptide or the polypeptide dimer. The plants are used for large-scale
CC production of mistletoe lectin for diagnostic or therapeutic purposes
CC (e.g. in cancer therapy). This sequence represents the mistletoe lectin
CC used in the method of the invention.
XX
SQ Sequence 564 AA;

Query Match 89.1%; Score 1269.5; DB 20; Length 564;
Best Local Similarity 88.9%; Pred. No. 3e-121;
Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;

QY 1 HQTGDEYFRFTLLRDYVSSGSFSNEIPLLKSTIPVSDAORFVLELTNOGSDSTAA 60
DB 43 HQTGDEYFRFTLLRDYVSSGSFSNEIPLLKSTIPVSDAORFVLELTNOGSDSTAA 102
QY 61 IDVTNLYVAVQAGDOSYFRLDAPDGAERHLFTGTTSSLPFGSYDLERRYAGHRQIP 120
DB 103 IDVTNLYVAVQAGDOSYFRLDAPDGAERHLFTGTTSSLPFGSYDLERRYAGHRQIP 162
QY 121 LGIELIQSVASALRYPGSGSTRAQARSLILLIOMISEAARENPJFWRARQYINGSGESFLPD 180
DB 163 LGIDQLIOSVYALRFPGSGSTRTQARSILLIOMISEAARENPJLMRRARQYINGSGESFLPD 222
QY 181 MYMELETSMGQOSTQYQOSTDGFENNFRLGISTGNFVLSNRYIASLAIMLEPYCRD 240
DB 223 VYMELETSMGQOSTQYQOSTDGFENNFRLGISTGNFVLSNRYIASLAIMLEPYCRD 282
QY 241 RPSSTDVRYWPIVIRPVLENSGAVDDVTCASEPTVRIY 279
DB 283 RPSSEVRYWPIVIRPVI-----ADVTCSASEPTVRIY 316

RESULT 6
AAV25970
ID AAV25970 standard; protein; 533 AA.
XX
XX AAV25970;
AC
XX 18-OCT-1999 (first entry)
DT
XX
XX Mistletoe lectin protein consensus sequence 1.
DE
XX
XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MIA; immunity;
KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KW cancer; cytotoxicity; antigen; isoform.
XX
OS Viscum album.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 15 /label= Asp, Glu
FT

FT Misc-difference 63 /label= Gly, Glu
FT Misc-difference 66 /label= Ile, Val
FT Misc-difference 75 /label= Leu, Ala
FT Misc-difference 107 /label= Asp, Arg or none
FT Misc-difference 113 /label= Asn, Thr
FT Misc-difference 117 /label= Pro, Thr
FT Misc-difference 134 /label= Asp, Glu
FT Misc-difference 141 /label= Ser, Thr
FT Misc-difference 145 /label= Phe, Tyr
FT Misc-difference 152 /label= Thr, Ala
FT Misc-difference 177 /label= Ala, Tyr
FT Misc-difference 180 /label= Tyr, Asp
FT Misc-difference 185 /label= Ala, Glu
FT Misc-difference 191 /label= Val, Met
FT Misc-difference 219 /label= Ile, Phe
FT Misc-difference 224 /label= Pro, Ser
FT Misc-difference 225 /label= Pro, Thr
FT Misc-difference 232 /label= Thr, Ser
FT Misc-difference 236 /label= Asp, Ser
FT Misc-difference 287 /label= Asn, Ser
FT Misc-difference 290 /label= Cys, Arg
FT Misc-difference 325 /label= Gly, Asn
FT Misc-difference 364 /label= Gly, Asp
FT Misc-difference 426 /label= Gly, Gln
FT Misc-difference 435 /label= Val, Asp
FT Misc-difference 439 /label= Gln, Lys
FT Misc-difference 442 /label= Gly or none
FT Misc-difference 443 /label= Arg, Lys
FT Misc-difference 464 /label= Cys, Ser, Val
FT Misc-difference 480 /label= Ala, Gly
FT Misc-difference 481 /label= Gly, Ala
FT Misc-difference 483 /label= Ser, Gly
FT Misc-difference 484 /label= Gly, Ser
FT Misc-difference 493 /label= Gly, Tyr
FT Misc-difference 500 /label= Asn, Ser, Thr, Lys
FT Misc-difference 501 /label= Ser, Gly
FT Misc-difference 502


```

FT      /label= Cys, Ser, Val
FT      Misc-difference 480
FT      /label= Ala, Gly
FT      Misc-difference 481
FT      /label= Gly, Ala
FT      Misc-difference 483
FT      /label= Ser, Gly
FT      Misc-difference 484
FT      /label= Gly, Ser
FT      Misc-difference 493
FT      /label= Gly, Tyr
FT      Misc-difference 500
FT      /label= Asn, Ser, Thr, Lys
FT      Misc-difference 501
FT      /label= Ser, Gly
FT      Misc-difference 502
FT      /label= Leu, Pro
FT      Misc-difference 503
FT      /label= Ala, Met
FT      Misc-difference 504
FT      /label= Met, Val
FT      Misc-difference 533
FT      /label= Pro, Phe
XX      DE19804210-A1.
XX      12-AUG-1999.
XX      03-FEB-1998; 98DE-1004210.
XX      03-FEB-1998; 98DE-1004210.
XX      (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX      PI Morris P, Stiefel T, Voelter W, Welters P;
XX      WPI; 1999-445335/38.
XX      Preparation of mistletoe lectins in heterologous systems,
XX      particularly for use as anticancer agents and immunostimulants
XX      Claim 4; Page 28-29; 78pp; German.
XX      This invention describes a novel mistletoe lectin (I) and its fragments
XX      which have antitumour and immunostimulatory activity. The A-chain (MLA)
XX      of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX      ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX      lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX      fragments are used to treat uncontrolled cell growth (particularly
XX      cancers) and if they lack cytotoxicity, to increase the strength of the
XX      immune response, particularly to a co-administered antigen
XX      (tumour-associated, bacterial or viral). The method allows production of
XX      mistletoe lectin, and its individual chains, in many different isoforms
XX      and on a large scale, at any time of the year. Recombinant products are
XX      free from toxins present in natural mistletoe extracts. This sequence
XX      represents a consensus sequence of the mistletoe lectin described in the
XX      specification.
SQ      Sequence 533 AA:
Query Match 85.0%; Score 1211; DB 20; Length 533;
Best Local Similarity 86.4%; Pred. No. 2.8e-115;
Matches 242; Conservative 6; Mismatches 26; Indels 6; Gaps 2;
OY      1 HOTGDEYFRFTILRLRYVSSGSNSRIPILRSTIPVSAQRVVELTNGGDSITAA 60
DB      10 HOTGAXYFRFTILRLRYVSSGSNSRIPILRSTIPVSAQRVVELTNGGDSITAA 69
OY      61 IDVTNLVVAQAGDOSYFLRDAAPDGAERHLFTGTPR-SSLPFGSYTDLERVAGHRDOI 119
DB      70 IDVTNXXVVAQAGDOSYFLRDAAPDGAERHLFTGTPR-SSLPFGSYTDLERVAGHRDOI 129
OY      120 PLGIEELIQSVASALRYPGSGSTRAQARSLIILQIMISEARFNPIFWRARQYINSGSEFLP 179

```

```

DB      130 PLGIXQLIQSVAXLRKRGSGSTRQARSLIILQIMISEARFNPILRKXQXINSGSEFLP 189
OY      180 DMTWLELETSWGOOSTOVQOSTGCVFNPPRLGISTGNFVTLSNVRDVIASLAIMLFVCR 239
DB      190 DXYWLELETSWGOOSTOVQOSTGCVFNPPRLGISTGNFVTLXNVXVIASLAIMLFVCG 249
OY      240 DRPSSDPVRYWPLVIRPVLENLSGAVDDVYCTASEPVRIV 279
DB      250 ERPSSDPVRYWPLVIRPV-----ADVYCSASEPVRIV 284

RESULT 8
AAV25976
ID      AAV25976 standard; protein; 533 AA.
AC      AAV25976;
DT      18-OCT-1999 (first entry)
XX      Mistletoe lectin protein consensus sequence 3.
XX      KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
XX      KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
XX      KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX      KW cancer; cytotoxicity; antigen; isoform.
XX      OS Viscum album.
XX      FH Key
XX      FT Misc-difference 15
XX      FT /label= Asp, Glu
XX      FT Misc-difference 63
XX      FT /label= Gly, Glu
XX      FT Misc-difference 66
XX      FT /label= Ile, Val
XX      FT Misc-difference 75
XX      FT /label= Leu, Ala
XX      FT Misc-difference 107
XX      FT /label= Asp, none
XX      FT Misc-difference 113
XX      FT /label= Asn, Thr
XX      FT Misc-difference 117
XX      FT /label= Pro, Thr
XX      FT Misc-difference 134
XX      FT /label= Asp, Glu
XX      FT Misc-difference 141
XX      FT /label= Ser, Thr
XX      FT Misc-difference 145
XX      FT /label= Phe, Tyr
XX      FT Misc-difference 152
XX      FT /label= Thr, Ala
XX      FT Misc-difference 177
XX      FT /label= Ala, Tyr
XX      FT Misc-difference 180
XX      FT /label= Tyr, Asp
XX      FT Misc-difference 185
XX      FT /label= Ala, Glu
XX      FT Misc-difference 191
XX      FT /label= Val, Met
XX      FT Misc-difference 219
XX      FT /label= Ile, Phe
XX      FT Misc-difference 224
XX      FT /label= Pro, Ser
XX      FT Misc-difference 225
XX      FT /label= Pro, Thr
XX      FT Misc-difference 232
XX      FT /label= Thr, Ser
XX      FT Misc-difference 236
XX      FT /label= Asp, Ser
XX      FT Misc-difference 287
XX      FT /label= Asn, Ser
XX      FT Misc-difference 290

```


FT /label- Cys, Arg
 FT Misc-difference 325
 FT /label- Gly, Asn
 FT Misc-difference 364
 FT /label- Gly, Asp
 FT Misc-difference 426
 FT /label- Gly, Gln
 FT Misc-difference 435
 FT /label- Val, Asp
 FT Misc-difference 439
 FT /label- Gln, Lys
 FT Misc-difference 442
 FT /label- Gly, none
 FT Misc-difference 443
 FT /label- Arg, Lys
 FT Misc-difference 464
 FT /label- Cys, Ser, Val
 FT Misc-difference 480
 FT /label- Ala, Gly
 FT Misc-difference 481
 FT /label- Gly, Ala
 FT Misc-difference 483
 FT /label- Ser, Gly
 FT Misc-difference 484
 FT /label- Gly, Ser
 FT Misc-difference 493
 FT /label- Gly, Tyr
 FT Misc-difference 500
 FT /label- Asn, Ser, Thr, Lys
 FT Misc-difference 501
 FT /label- Ser, Gly
 FT Misc-difference 502
 FT /label- Leu, Pro
 FT Misc-difference 503
 FT /label- Ala, Met
 FT Misc-difference 504
 FT /label- Met, Val
 FT Misc-difference 533
 FT /label- Pro, Phe
 XX DE19804210-A1.
 XX
 XX 12-AUG-1999.
 XX
 XX 03-FEB-1998; 98DE-1004210.
 XX
 XX 03-FEB-1998; 98DE-1004210.
 XX
 XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 XX PI Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI; 1999-445335/38.
 XX
 XX Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 XX Claim 40; Page 37-38; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of the mistletoe lectin described in the
 CC specification.

XX
 SO Sequence 533 AA;
 Query Match 85.0%; Score 1211; DB 20; Length 533;
 Best Local Similarity 86.4%; Pred. No. 2,8e-115;
 Matches 242; Conservative 6; Mismatches 26; Indels 6; Gaps 2;
 QY 1 HOTGDEYFRITLLRDVSSGFSNRPILRSTIPVSDORFVLELNOGDSITAA 60
 DB 10 HOTGDEXFRITLLRDVSSGFSNRPILRSTIPVSDORFVLELNOGDSITAA 69
 QY 61 IDVTNLYVVAQAGDOSYFLRDAPGAEHLFTGTR-SSLPTGYSTDLERYAGHDDQI 119
 DB 70 IDVTNXXVVAQAGDOSYFLRDAPGAEHLFTGTRXSSLPEFGSYXDLERYAGHDDQI 129
 QY 120 PLGIEELIQSVSALRYPGSGTRFAQARSLIILQMISEARNP1PFWARQYINSGESFLP 179
 DB 130 PLGIXQLIQSVXALRXPGSGTRFAQARSLIILQMISEARNP1LWXRXVYIASLAIMLFCVCG 189
 QY 180 DMYMLETSMGQOSTOVQSTDGVPNNPRLIGISTGFVTLNVRDVIASLAIMLFCVCR 239
 DB 190 DXYMLETSMGQOSTOVQSTDGVPNNPRLAIXXGFEVTLXNVXVYIASLAIMLFCVCG 249
 QY 240 DRPSSDVRYWPLYRPLYLENSGAVDDVCTPASEPTVRIIV 279
 DB 250 ERPSSDVRYWPLYRPLYI-----ADDVTCASSEPVRIV 284
 RESULT 9
 AAB47090
 ID AAB47090 standard; Protein: 254 AA.
 XX
 XX AAB47090;
 XX
 XX 16-MAY-2001 (first entry)
 XX
 XX A-chain isoform for biosynthesis of a Korean mistletoe lectin #1.
 XX
 XX Isoform: A-chain; B-chain: biosynthesis; lectin; Korean mistletoe;
 KW KML; tumour; KM-110; KML-C; KML-BP; KML-11U; KML-11L;
 KW heparin binding protein.
 XX
 XX Viscum album coloratum.
 XX
 XX EP1074560-A2.
 XX
 XX 07-FEB-2001.
 XX
 XX 27-JUL-2000; 2000EP-0402168.
 XX
 XX 27-JUL-1999; 99KR-0030638.
 XX
 XX (MIST-) MISTLE BIOTECH CO LTD.
 XX
 XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 PI Kang T, Park C;
 XX
 XX WPI; 2001-171044/18.
 DR N-PSDB; AAC85472.
 XX
 PT Novel lectin proteins isolated from Korean mistletoe, useful for
 PT enhancing immunity and effectuating anti-tumoral activity -
 XX
 PS Claim 2; Page 25-26; 62pp; English.
 XX
 CC The sequences given in AAB47090-92 are isoforms of an A-chain gene
 CC involved in biosynthesis of lectins isolated from Korean mistletoe.
 CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
 CC for treating tumours. The KML's are isolated from a protein fraction
 CC derived from the leaves, stems and fruits of Korean mistletoe, which
 CC is designated KM-110. One of the isolates, KML-C, was shown to be
 CC effective against colon 26-M3.1 carcinoma and L5178Y-MJ25 lymphoma.
 CC

SQ Sequence 254 AA;
 Query Match 81.1%; Score 1155; DB 22; Length 254;
 Best Local Similarity 93.1%; Pred. No. 5.3e-110;
 Matches 220; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 HQTGDEYFRFTLLRDYVSSGFSFSENEIPLRQSTIPVSDAORFVLELNOGGDSITAA 60
 Db 10 HQTGDEYFRFTLLRDYVSSGFSFSENEIPLRQSTIPVSDAORFVLELNOGGDSITAA 69
 QY 61 IDVTNLYVVAOAGDOSYFLRDAPDGAERHLFTGTTSSSLPFTGSYTDLERVAGHRDQ 120
 Db 70 IDVTNLYVVAOAGDOSYFLRDAPDGAERHLFTGTTSSSLPFTGSYTDLERVAGHRDQ 129
 QY 121 LGIEELIOSVSLRYPGSGSTRPAORSLILILOMISEAARNPIFWRARQYINGGESFLP 180
 Db 130 LGRELIQSVALRFPGSNTRPAORSLILILOMISEAARNPILMRRQYISSGGSFLPD 189
 QY 181 MYMLEETSMWGQSTQVOQSTDGVEFNNPFRIGISTGNFVTLNVRDYIASLAIMLFVC 240
 Db 190 TYILQLETSWGQSTQVOHSTDGVEFNNPFRILITISTGVFVTLNVRDYIASLAIMLFVC 249
 QY 241 RPSSS 245
 Db 250 RPSSS 254

RESULT 10
 AAY25981
 ID AAY25981 standard; Protein; 256 AA.
 AC AAY25981;
 DT 18-OCT-1999 (first entry)
 DE Mistletoe lectin A2 protein fragment.
 XX
 KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A2.
 OS Viscum album.
 XX
 PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI; 1999-445335/38.
 DR N-PSDB; AA209105.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 PS Claim 8; Fig 3b; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of

CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of the mistletoe lectin A2 protein.
 CC
 SQ Sequence 256 AA;
 Query Match 80.8%; Score 1152; DB 20; Length 256;
 Best Local Similarity 93.1%; Pred. No. 1.1e-109;
 Matches 230; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 HQTGDEYFRFTLLRDYVSSGFSFSENEIPLRQSTIPVSDAORFVLELNOGGDSITAA 60
 Db 10 HQTGDEYFRFTLLRDYVSSGFSFSENEIPLRQSTIPVSDAORFVLELNOGGDSITAA 69
 QY 61 IDVTNLYVVAOAGDOSYFLRDAPDGAERHLFTGTT--RSSLPFTGSYTDLERVAGHRDQ 118
 Db 70 IDVTNLYVVAOAGDOSYFLRDAPDGAERHLFTGTTTRPSSSLPFTGSYTDLERVAGHRDQ 129
 QY 119 IPIGIEELIOSVSLRYPGSGSTRPAORSLILILOMISEAARNPIFWRARQYINGGESFL 178
 Db 130 IPIGIEELIOSVSLRYPGSGSTRPAORSLILILOMISEAARNPILMRRQYISSGGSFL 189
 QY 179 PDMYMLEETSMWGQSTQVOQSTDGVEFNNPFRIGISTGNFVTLNVRDYIASLAIMLFVC 238
 Db 190 PDMYMLEETSMWGQSTQVOHSTDGVEFNNPFRILITISTGVFVTLNVRDYIASLAIMLFVC 249
 QY 239 RDRPSSS 245
 Db 250 GERPSSS 256

RESULT 11
 AAY25984
 ID AAY25984 standard; Protein; 256 AA.
 AC AAY25984;
 DT 18-OCT-1999 (first entry)
 DE Mistletoe lectin A2 (variant) protein fragment.
 XX
 KW Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A2.
 OS Viscum album.
 XX
 PN DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 03-FEB-1998; 98DE-1004210.
 XX
 PR 03-FEB-1998; 98DE-1004210.
 XX
 PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
 XX
 PI Morris P, Stiefel T, Voelter W, Welters P;
 XX WPI; 1999-445335/38.
 DR N-PSDB; AA209108.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 XX particularly for use as anticancer agents and immunostimulants
 PS Disclosure; Fig 6b; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and

CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a fragment of a mistletoe lectin A2 protein variant.

XX Sequence 256 AA:

Query Match 80.8%; Score 1152; DB 20; Length 256;

Best Local Similarity 93.1%; Pred. No. 1.1e-109;

Matches 220; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 HQTGDEYFRFTLLRDYSSGSFSNEIPLLRQSTIPVSDAQRFLVELTNGGDSITAA 60

DB 10 HQTGDEYFRFTLLRDYSSGSFSNEIPLLRQSTIPVSDAQRFLVELTNGGDSITAA 69

QY 61 IDVTNLVYVAYQAGDSYFLRDAPDGAERHLEFTGTT--RSLEPFGSYDLERYAGHRDQ 118

DB 70 IDVTNLVYVAYQAGDSYFLRDAPRGAERHLEFTGTTDRSSLPFGSYDLERYAGHRDQ 129

QY 119 IPLGIEELIQSVSALRYPGSGSTRAQARSLLILQIMISEARFNPILMRARQYINGSGESFL 178

DB 130 IPLGIEELIQSVSALRYPGSGSTRAQARSLLILQIMISEARFNPILMRARQYINGSGESFL 189

QY 179 PDWYMLEFTSWGQOSTGVQVSTGVNPNPRLGISTGNFVTLISNVDYIASLAIMLFVC 238

DB 190 PDWYMLEFTSWGQOSTGVQVSTGVNPNPRLAISTGNFVTLISNVDYIASLAIMLFVC 249

QY 239 RDRPSSS 245

DB 250 GERPSSS 256

RESULT 12

AAW64661 ID AAW64661 standard; Protein; 252 AA.

XX AAW64661;

DT 23-OCT-1998 (first entry)

XX Mistletoe rMLA variant protein.

XX Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;

KW intracellular; processing module; protease recognition; targeting module;

KM internalisation; treatment; disorder; cell proliferation; activation;

XX autoimmune disease; allergy; tumour; ricin; translocation; ss.

OS Viscum album.

XX Key Location/Qualifiers

FT Protein 1..252

FT /note="partial"

XX W09829540-A2.

XX 09-JUL-1998.

XX 02-JAN-1998; 98WO-EP00009.

XX 02-JAN-1997; 97EP-0100012.

XX (BRA1-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.

XX Eck J, Schmidt A, Zinke H;

XX WPI: 1998-388122/33.

XX N-PSDB: AAV51343.

PT Nucleic acid encoding fusion protein containing mistletoe lectin A
 PI chain - useful for treatment of proliferative and autoimmune
 XX diseases, allergies and tumours

XX Disclosure; Fig 11a'; 115pp; German.

This sequence encodes a variant mistletoe lectin A-chain, rMLA. This
 CC sequence can be used in the construction of a fusion protein which
 CC comprises an effector module that is cytotoxic intracellularly, a
 CC processing module covalently bonded to the effector module and
 CC containing a protease recognition sequence, and a targeting module
 CC covalently bonded to the processing module, able to bind specifically to
 CC the surface of a cell so as to mediate internalisation of the fusion
 CC protein. Such a fusion protein can be used for treating disorders
 CC involving proliferation and/or elevated activation of cells, especially
 CC autoimmune disease, allergy and tumours. The proteins can be administered
 CC at 1 ng to 500 mg g/kg/day, or for ex vivo use at 1 pg to 500 ng/mL.
 CC Fusion proteins can develop toxic activity in a wide range of target
 CC cells. The processing module prevents extracellular dissociation, and
 CC fusion proteins based on mistletoe lectin A-chain are far more active
 CC than those based on ricin and do have the associated problems of
 CC non-specific toxicity. The protein may be expressed in a non-glycosylated
 CC form that does not bind to sugar receptors in the liver, and which has a
 CC long half-life in the blood. Where the mistletoe lectin B-chain is used,
 CC it actively assists in translocation of the ML A-chain from the
 CC endoplasmic reticulum to the cytoplasm.

XX Sequence 252 AA:

Query Match 79.4%; Score 1131; DB 19; Length 252;

Best Local Similarity 90.5%; Pred. No. 1.5e-107;

Matches 220; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 HQTGDEYFRFTLLRDYSSGSFSNEIPLLRQSTIPVSDAQRFLVELTNGGDSITAA 60

DB 10 HQTGDEYFRFTLLRDYSSGSFSNEIPLLRQSTIPVSDAQRFLVELTNGGDSITAA 69

QY 61 IDVTNLVYVAYQAGDSYFLRDAPDGAERHLEFTGTTSSLPFGSYDLERYAGHRDQIP 120

DB 70 IDVTNLVYVAYQAGDSYFLRDAPRGAERHLEFTGTTSSLPFGSYDLERYAGHRDQIP 129

QY 121 IGLIEELIQSVSALRYPGSGSTRAQARSLLILQIMISEARFNPILMRARQYINGSGESFLPD 180

DB 130 IGLIEELIQSVSALRYPGSGSTRAQARSLLILQIMISEARFNPILMRARQYINGSGESFLPD 189

QY 181 MYWLEFTSWGQOSTGVQVSTGVNPNPRLGISTGNFVTLISNVDYIASLAIMLFVC 240

DB 190 MYWLEFTSWGQOSTGVQVSTGVNPNPRLAISTGNFVTLISNVDYIASLAIMLFVC 249

QY 241 RPS 243

DB 250 RPS 252

RESULT 13

AAW10022 ID AAW10022 standard; Protein; 253 AA.

XX AAW10022;

DT 18-DEC-1997 (first entry)

XX Prepro mistletoe lectin A chain.

XX Mistletoe; lectin; cytotoxic; A chain; B chain; dimer.

XX Viscum album.

XX EP751221-A1.

XX 02-JAN-1997.

PF 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 95EP-0109949.
 XX
 PA (MADU) MADDAUS KOELN AG.
 XX
 PI Baur A, Eck J, Lentzen H, Zinke H;
 XX WPI; 1997-054678/06.
 DR N-PSDB; AAT70474.
 XX
 PT Nucleic acid encoding pre-pro form of mistletoe lectin - for
 PT therapeutic or diagnostic use
 XX
 PS Claim 12; Fig 4A; 30pp; German.
 CC Mistletoe lectin is a cytotoxic agent that has been used for tumour
 CC therapy. It can be used in immunotoxins and medicaments. Nucleic
 CC acid fragments can be used in diagnostic methods. Mistletoe lectin (
 CC AAT70473) comprises an A chain (AAT70474) and a B chain (AAT70475).
 CC
 SQ Sequence 253 AA;

Query Match 79.4%; Score 1131; DB 18; Length 253;
 Best Local Similarity 90.5%; Pred. No. 1.5e-107;
 Matches 220; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 HQTGDEYFRFTLLRDYVSSGFSFNEIPILRQSTIPVSDAQRVLYVELTNGGDSITAA 60
 |||||
 DB 11 HQTGDEYFRFTLLRDYVSSGFSFNEIPILRQSTIPVSDAQRVLYVELTNGGDSITAA 70
 QY 61 IDVTNLYVVAQAGDOSYFLRDAPDGAERHLEFTGTRSSLPFGTSYDLERYAGHRDIP 120
 |||||
 DB 71 IDVTNLYVVAQAGDOSYFLRDAPDGAERHLEFTGTRSSLPFGTSYDLERYAGHRDIP 130
 QY 121 LGIEELIQSVASALRYPGGSTRQAARSLITLIOMISEARFNPIMRAKQYINSGSFLPD 180
 |||||
 DB 131 LGIDQLIQSVATLRFPGGSTRQAARSLITLIOMISEARFNPIMRAKQYINSGSFLPD 190
 QY 181 MYMLELETSMGQOSTVOQSTGVFNPNPRLIGISTGNFVTLISNVDYIASLAIMLFVCRD 240
 :|||||
 DB 191 VYMLELETSMGQOSTVOQSTGVFNPNPRLIAPGNFVTLISNVDYIASLAIMLFVCGE 250
 QY 241 RPS 243
 |||
 DB 251 RPS 253

RESULT 14
 AAM90125
 ID AAM90125 standard; Protein; 253 AA.
 XX

AC AAM90125;
 DT 30-APR-1999 (first entry)
 XX

DE Mistletoe ML A-chain protein.
 XX

KW MI; mistletoe; lectin; MLA; A-chain; transgenic plant; glycosylation;
 KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;
 KW cancer.
 XX

OS Viscum album.
 XX

PN EP884388-A1.
 XX

PD 16-DEC-1998.
 XX

PF 26-JUN-1995; 95EP-0109949.
 XX

PR 26-JUN-1995; 95EP-0109949.
 PR 26-JUN-1995; 98EP-0105660.
 XX

PA (MADU) MADDAUS KOELN AG.

PI Baur A, Eck J, Lentzen H, Zinke H;

DR WPI; 1999-026582/03.

DR N-PSDB; AAV74180.

PT New transgenic plant expressing mistletoe lectin - useful for
 PT producing recombinant lectin in e.g. cancer diagnosis and therapy

PS Disclosure; Fig 4a; 30pp; German.

This invention describes a novel transgenic plant transformed with a
 CC vector capable of encoding a mistletoe (Viscum album) lectin
 CC preproprotein or a biologically active fragment. The specification
 CC also describes a polypeptide produced by a plant where the polypeptide
 CC exhibits at least one enzymatic modification other than the glycosylation
 CC that occurs in Viscum album or the polypeptide is a fusion protein, a
 CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the
 CC polypeptide or the polypeptide dimer. The plants are used for large-scale
 CC production of mistletoe lectin for diagnostic or therapeutic purposes
 CC (e.g. in cancer therapy). This sequence represents the mistletoe lectin
 CC A-chain which is contained in expression vector pT7MLA.
 CC
 SQ Sequence 253 AA;

Query Match 79.4%; Score 1131; DB 20; Length 253;
 Best Local Similarity 90.5%; Pred. No. 1.5e-107;
 Matches 220; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 HQTGDEYFRFTLLRDYVSSGFSFNEIPILRQSTIPVSDAQRVLYVELTNGGDSITAA 60
 |||||
 DB 11 HQTGDEYFRFTLLRDYVSSGFSFNEIPILRQSTIPVSDAQRVLYVELTNGGDSITAA 70
 QY 61 IDVTNLYVVAQAGDOSYFLRDAPDGAERHLEFTGTRSSLPFGTSYDLERYAGHRDIP 120
 |||||
 DB 71 IDVTNLYVVAQAGDOSYFLRDAPDGAERHLEFTGTRSSLPFGTSYDLERYAGHRDIP 130
 QY 121 LGIEELIQSVASALRYPGGSTRQAARSLITLIOMISEARFNPIMRAKQYINSGSFLPD 180
 |||||
 DB 131 LGIDQLIQSVATLRFPGGSTRQAARSLITLIOMISEARFNPIMRAKQYINSGSFLPD 190
 QY 181 MYMLELETSMGQOSTVOQSTGVFNPNPRLIGISTGNFVTLISNVDYIASLAIMLFVCRD 240
 :|||||
 DB 191 VYMLELETSMGQOSTVOQSTGVFNPNPRLIAPGNFVTLISNVDYIASLAIMLFVCGE 250
 QY 241 RPS 243
 |||
 DB 251 RPS 253

RESULT 15
 AAM64659
 ID AAM64659 standard; Protein; 252 AA.
 XX

AC AAM64659;
 DT 23-OCT-1998 (first entry)
 XX

DE Mistletoe rMLA protein.
 XX

KW Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; ricin; translocation.
 XX

OS Viscum album.
 XX

PN WO9829540-A2.
 XX

PD 09-JUL-1998.
 XX

PF 02-JAN-1998; 98WO-EP00009.
 XX

```
XX 02-JAN-1997; 97EP-0100012.
PR (BRA1-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
XX
XX Eck J, Schmidt A, Zinke H;
XX
XX WPI; 1998-388122/33.
DR N-PSDB; AAV51341.
XX
PR Nucleic acid encoding fusion protein containing mistletoe lectin A
PR chain - useful for treatment of proliferative and autoimmune
PR diseases, allergies and tumours
XX
XX Disclosure: Fig 11a; 115pp; German.
XX
XX This sequence represents a lectin A-chain, rMLA, isolated from mistletoe.
XX This sequence can be used in the construction of a fusion protein which
XX comprises an effector module that is cytotoxic intracellularly, a
XX processing module covalently bonded to the effector module and
XX containing a protease recognition sequence, and a targeting module
XX covalently bonded to the processing module, able to bind specifically to
XX the surface of a cell so as to mediate internalisation of the fusion
XX protein. Such a fusion protein can be used for treating disorders
XX involving proliferation and/or elevated activation of cells, especially
XX autoimmune disease, allergy and tumours. The proteins can be administered
XX e.g. by injection or topically but especially by intravenous injection,
XX at 1 ng to 500 nm g/kg/day, or for ex vivo use at 1 pg to 500 ng/mL.
XX Fusion proteins can develop toxic activity in a wide range of target
XX cells. The processing module prevents extracellular dissociation, and
XX fusion proteins based on mistletoe lectin A-chain are far more active
XX than those based on ricin and do have the associated problems of
XX non-specific toxicity. The protein may be expressed in a non-glycosylated
XX form that does not bind to sugar receptors in the liver, and which has a
XX long half-life in the blood. Where the mistletoe lectin B-chain is used,
XX it actively assists in translocation of the ML A-chain from the
XX endoplasmic reticulum to the cytoplasm.
XX
SQ Sequence 252 AA;
Query Match 79.1%; Score 1127; DB 19; Length 252;
Best Local Similarity 90.5%; Pred. No. 3.9e-107;
Matches 219; Conservative 10; Mismatches 13; Indels 0; Gaps 0;
QY 1 HQTGDEFFRFTLLRDVYSSSEFSENEIPLRQSTIPYSDAQRFLVELTNOGDSITAA 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
11 HQTGEYFRFTLLRDYSSGSFSNEIPLRQSTIPYSDAQRFLVELTNOGDSITAA 70
QY 61 IDVTNLVYVAYOAGDQSYFLRDAPGAEERHLEFTGRSSLPPTGSYTDLERYAGHRDIP 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
71 IDVTNLVYVAYOAGDQSYFLRDAPGAEERHLEFTGRSSLPPTGSYTDLERYAGHRDIP 130
QY 121 LGIEELIQSVSLARYPGSSTRAQARSLLIQLMISEARFNPIFWARQYINSGESFLPD 180
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
131 LGIDQLIQSVTLRFPGGSTRTQARSILLIQLMISEARFNPIELWRARQYINSGESFLPD 190
QY 181 MYMLELFTSMGQSTQVOQSTDGVFNPNPRLGISTGNFVTLSNVDVYASLAIMLFVCRD 240
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
191 VYMLELFTSMGQSTQVOQSTDGVFNPNPRLAIPPGNFVTLTNVDVYASLAIMLFVCGE 250
QY 241 RP 242
DB ||
251 RP 252
```

Search completed: March 18, 2003, 08:25:13
Job time : 52.0704 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:25:21 ; Search time 18.2261 Seconds
(without alignments)
450.397 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1425
Sequence: 1 HQTGDEYFRFTLLRDYSS.....NSGAVDDYCTASEPTVRIV 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6CTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269.5	89.1	564	4	US-08-776-059-35 Sequence 35, Appl
2	1131	79.4	253	4	US-08-776-059-31 Sequence 31, Appl
3	1088	76.4	235	4	US-08-776-059-39 Sequence 39, Appl
4	466.5	32.7	250	1	US-08-378-761A-71 Sequence 71, Appl
5	466.5	32.7	250	1	US-08-485-286-71 Sequence 71, Appl
6	423	29.7	540	1	US-08-378-761A-77 Sequence 77, Appl
7	423	29.7	540	1	US-08-485-286-77 Sequence 77, Appl
8	408.5	28.7	534	2	US-08-356-786-10 Sequence 10, Appl
9	383	26.9	267	1	US-07-901-707-1 Sequence 1, Appl
10	383	26.9	267	1	US-07-988-430-1 Sequence 1, Appl
11	383	26.9	267	1	US-08-218-303-16 Sequence 16, Appl
12	383	26.9	267	1	US-08-425-336-1 Sequence 1, Appl
13	383	26.9	267	1	US-08-488-113B-1 Sequence 1, Appl
14	383	26.9	267	1	US-08-477-484B-1 Sequence 1, Appl
15	383	26.9	267	2	US-08-646-360-1 Sequence 1, Appl
16	383	26.9	267	2	US-08-338-793D-61 Sequence 61, Appl
17	383	26.9	267	4	US-08-839-765-1 Sequence 1, Appl
18	383	26.9	267	4	US-09-136-389-1 Sequence 1, Appl
19	383	26.9	267	4	US-09-610-838-1 Sequence 1, Appl
20	383	26.9	267	5	PCR-US92-09487-1 Sequence 1, Appl
21	383	26.9	268	2	US-08-356-786-8 Sequence 1, Appl
22	379	26.6	290	1	US-08-378-761A-27 Sequence 8, Appl
23	379	26.6	290	1	US-08-485-286-27 Sequence 27, Appl
24	379	26.6	290	1	US-08-324-301-15 Patent No. 5248606
25	354	24.8	282	1	US-07-901-707-4 Sequence 15, Appl
26	340.5	23.9	263	1	US-07-988-430-4 Sequence 4, Appl
27	340.5	23.9	263	1	US-07-988-430-4 Sequence 4, Appl

28	340.5	23.9	263	1	US-08-425-336-4 Sequence 4, Appl
29	340.5	23.9	263	1	US-08-488-113B-4 Sequence 4, Appl
30	340.5	23.9	263	1	US-08-477-484B-4 Sequence 4, Appl
31	340.5	23.9	263	2	US-08-646-360-4 Sequence 4, Appl
32	340.5	23.9	263	4	US-08-839-765-4 Sequence 4, Appl
33	340.5	23.9	263	4	US-09-136-389-4 Sequence 4, Appl
34	340.5	23.9	263	4	US-09-610-838-4 Sequence 4, Appl
35	340.5	23.9	263	5	PCR-US92-09487-4 Sequence 4, Appl
36	336.5	23.6	286	1	US-08-324-301-13 Sequence 13, Appl
37	336	23.6	267	1	US-08-378-761A-74 Sequence 74, Appl
38	336	23.6	267	1	US-08-485-286-74 Sequence 74, Appl
39	329	23.1	289	1	US-07-923-692C-4 Sequence 4, Appl
40	329	23.1	289	1	US-08-184-237-4 Sequence 4, Appl
41	329	23.1	289	2	US-08-482-920-4 Sequence 4, Appl
42	329	23.1	289	3	US-08-484-341-4 Sequence 4, Appl
43	329	23.1	289	4	US-08-483-502-4 Sequence 4, Appl
44	329	23.1	289	4	US-09-726-651A-4 Sequence 4, Appl
45	326	22.9	247	1	US-08-488-113B-6 Sequence 6, Appl

ALIGNMENTS

RESULT 1									
US-08-776-059-35									
Sequence 35, Application US/08776059B									
Patent No. 6271368									
GENERAL INFORMATION:									
APPLICANT: LENTZEN, Hans									
APPLICANT: ECK, Jurgen									
APPLICANT: BAUR, Axel									
APPLICANT: ZINKE, Holger									
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)									
FILE REFERENCE: 674503-2003									
CURRENT APPLICATION NUMBER: US/08/776, 059B									
CURRENT FILING DATE: 1999-06-19									
EARLIER APPLICATION NUMBER: PCT/EP96/02273									
EARLIER FILING DATE: 1996-06-25									
EARLIER APPLICATION NUMBER: 95109949.8									
EARLIER FILING DATE: 1995-06-26									
NUMBER OF SEQ ID NOS: 56									
SOFTWARE: PatentIn Ver. 2.0									
SEQ ID NO 35									
LENGTH: 564									
TYPE: PRT									
ORGANISM: Viscum album									
US-08-776-059-35									
Query Match									
Best local similarity 89.1% Score 1269.5; DB 4; Length 564;									
Matches 248; Conservative 12; Mismatches 14; Indels 5; Gaps 1;									
QY	1	HQTGDEYFRFTLLRDYSSGFSFNEIPLRKQSTIPVSDAQRFLVLTJNOGDSITAA	60						
DB	43	HQTGDEYFRFTLLRDYSSGFSFNEIPLRKQSTIPVSDAQRFLVLTJNOGDSITAA	102						
QY	61	IDVTMLVYVAAQAGQSTFLRLAPGARGRLFTGTTRSSLPFTGSDYTDLERRARRDIP	120						
DB	103	IDVTMLVYVAAQAGQSTFLRLAPGARGRLFTGTTRSSLPFTGSDYTDLERRARRDIP	162						
QY	121	IGIERLISVSALRRPGGSTRQAARSLIILQIMSEARFPDIFPBARQVYNSGESFLPD	180						
DB	163	IGIDLDLISVTLRRPGGSTRQAARSLIILQIMSEARFPDIFPBARQVYNSGESFLPD	222						
QY	181	KYMLELTSWGQOSTVOVQSTDGVFNNPRLGISTGNFVLTJNVRDYIASLAIMLFCYCD	240						
DB	223	YMLELTSWGQOSTVOVQSTDGVFNNPRLGISTGNFVLTJNVRDYIASLAIMLFCYCD	282						
QY	241	RSSSDVRYWPLVTRPVLENGAADDVYCTASEPTVRIV	279						
DB	283	RSSSEVRYWPLVTRPVLENGAADDVYCTASEPTVRIV	316						

```

RESULT 2
US-08-776-059-31
Sequence 31, Application US/08776059B
Patent No. 6271368
GENERAL INFORMATION:
APPLICANT: LENTZEN, Hans
APPLICANT: ECK, Jurgen
APPLICANT: BAUER, Axel
APPLICANT: ZINKE, Holger
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
FILE REFERENCE: 674503-2003
CURRENT APPLICATION NUMBER: US/08/776, 059B
CURRENT FILING DATE: 1999-06-19
EARLIER APPLICATION NUMBER: PCT/EP96/02273
EARLIER FILING DATE: 1996-06-25
EARLIER APPLICATION NUMBER: 9510994.9
EARLIER FILING DATE: 1995-06-26
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 31
LENGTH: 253
TYPE: prt
ORGANISM: Viscum album
US-08-776-059-31

```

Query Match	79.4%;	Score 1131;	DB 4;	Length 253;
Best Local Similarity	90.5%;	Pred. No. 2.7e-112;		
Matches 220;	Conservative 10;	Mismatches 13;	Indels 0;	Gaps 0;

Qy	1	HOTTGDEEERFTLLRDVSSGSGSNELPILROSTIVPSAQRVVELNOCGDSITTA	60
Qy	1	HOTTGDEEERFTLLRDVSSGSGSNELPILROSTIVPSAQRVVELNOCGDSITTA	60
Db	11	HOTTGEEERFTLLRDVSSGSGSNELPILROSTIVPSAQRVVELNOCGDSITTA	70
Qy	61	IDVTNLVYVAYOAGDQSYFLRDADGAEERHLEFTGTTRSSLEPFGSYTDLERYAGHRDQIP	120
Db	71	IDVTNLVYVAYOAGDQSYFLRDADGAEERHLEFTGTTRSSLEPFGSYTDLERYAGHRDQIP	130
Qy	121	LGIEERLOSVALRPGSGSTQARASLITLLOMISEARFNPILPWRARQYINSGESFLPD	180
Db	131	LGIDDLDSVTLARPGSGSTQARSLITLLOMISEARFNPILPWRARQYINSGESFLPD	190
Qy	181	MYMELLETSMGQOSTQOVQOSTDGVFNNEPFRLGISTGPFVLSNVRDIYASIALINLFCYCD	240
Db	191	MYMELLETSMGQOSTQOVQOSTDGVFNNEPFRLGISTGPFVLSNVRDIYASIALINLFCYCD	250
Qy	241	RPS 243	
Db	251	RPS 253	

```

? RESULT 3
? US-08-776-059-39
? Sequence 39, Application US/08776059B
? Patent No. 6271368
? GENERAL INFORMATION:
? APPLICANT: LENTZEN, Hans
? APPLICANT: ECK, Jurgen
? APPLICANT: BAUR, Axel
? APPLICANT: ZINKE, Holger
? TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
? FILE REFERENCE: 674503-2003
? CURRENT APPLICATION NUMBER: US/08/776,059B
? EARLIER FILING DATE: 1999-06-19
? EARLIER APPLICATION NUMBER: PCT/EP96/02273
? EARLIER FILING DATE: 1996-06-25
? EARLIER APPLICATION NUMBER: 95109949.8
? EARLIER FILING DATE: 1995-06-26
? NUMBER OF SEQ ID NOS: 56
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 39
? LENGTH: 235
? TYPE: PRT
? ORGANISM: Viscum album

```

US-08-776-059-39

Query Match	76.48;	Score 1088;	DB 4;	Length 235;
Best Local Similarity	90.68;	Pred. No. 9.2e-108;		
Matches 213; Conservative	9;	Mismatches 13;	Indels 0;	Gaps 0;

[illegible]

RESULT 4
US-08-378-761A-71
: Sequence 71, Application US/08378761A

```

1 Patent NO. 5635384
2 GENERAL INFORMATION:
3 APPLICANT: WALSH, TERENCE A
4 APPLICANT: HEY, TIMOTHY D
5 APPLICANT: MORAN, ALICE ER
6 TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
7 TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
8 TITLE OF INVENTION: USING
9 NUMBER OF SEQUENCES: 81
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: ANDREA T. BORUCKI
12 STREET: 9330 ZIONSVILLE ROAD
13 CITY: INDIANAPOLIS
14 STATE: IN
15 COUNTRY: US
16 ZIP: 46268
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: PatentIn Release #1.0, Version #1.25
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/378,761A
24 FILING DATE: 26-JAN-1995
25 CLASSIFICATION: 435
26 ATTORNEY/AGENT INFORMATION:
27 NAME: BORUCKI, ANDREA T
28 REGISTRATION NUMBER: 33651
29 REFERENCE/DOCKET NUMBER: 38272B
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (317) 337-4846
32 INFORMATION FOR SEQ ID NO: 71:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 250 amino acids
35 TYPE: amino acid
36 STRANDEDNESS: single
37 TOPOLOGY: linear
38 MOLECULE TYPE: protein
39 US-08-378-761A-71

```

	Query Match	Similarity	Score	DB	Length
Best	Local	108%	Conservative	36	Mismatches 78; Indels 25; Gaps 6;
Qy	4	TGDEYFRFTLLRDYVSSGSFNSFELPLLRQSTIPVSDAQRFLVVELLHNOGGDSTIAADY	63		
Db	13	TSQSKQCEIARREKL-NGGLRHDIPVDPDP-TLOERKRYTIVLPLNSDSEIEVGD	70		

OY	64	TNKKVVAQAQDSDSYFLRBDAPDGAERHLEFGCTRRSLPEFGSTYDLERYAGH-RDQIPIC	122
Dd	71	TNAVVAAFRAGTOSYFLRDAPSSASDYLFETGDOSH.LPFGYGDDLERMHQSROOPIIC	130
OY	123	IEELIQSVSALRYPGSGSTRAOARSLITLIIOMISEAFENFIEMRAROYINSGESPLPDWY	182
Dd	131	LQALTHGHSFFRSCGNONEEKARLLIYIIONVAFANFRIISNRVRSYIOGTAFOPDAA	190
OY	183	MLELETSMGQGSSTVOQGSDGVFNPNPRIGISTGNFTLSNVHD-----VIASLT	231
Dd	191	MISLENNW-DNLRGVQESVDTFPMQ-----VTLININNEPYIVDSLSHPYAVL	239
OY	232	AIMLFVC	238
Dd	240	AIMLFVC	246

```

US-08-485-286-71
: Sequence 71, Application US/08485286
: Patent No. 5646026
: Patent No. 5646026 5646119
: GENERAL INFORMATION:
: APPLICANT: WALSH, TERENCE A
: APPLICANT: HEY, TIMOTHY D
: TITLE OF INVENTION: MORGAN, ALICE ER
: TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
: TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
: NUMBER OF SEQUENCES: 81
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ANDREA T. BORUCKI
: STREET: 9330 ZIONSVILLE ROAD
: CITY: INDIANAPOLIS
: STATE: IN
: COUNTRY: US
: ZIP: 46268
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/485,286
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/378761
: FILING DATE: 26-JAN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: BORUCKI, ANDREA T
: REGISTRATION NUMBER: 33651
: REFERENCE/DOCKET NUMBER: 38272B
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (317) 337-4846
: INFORMATION FOR SEQ ID NO: 71:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 250 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: IS-08-485-286-71

```

```

Query Match: 32.7%; Score 466.5; DB 1, Length 250;
Best Match Similarity 43.7%; Pred. NO. 1.3e-41;
Matches 108; Conservative 36; Mismatches 78; Indels 25; Gaps 6;

QY      4 TGEYRFTLLRDYVSSGFSNEIPLIKOSTIPVSDQRFVELTNOGDSITAIADV 63
      | : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 13 TSGSTQFTEAKRERKRGGLHIDIPDPTFLQERNRIYVELNSNSTEISLEIGIV 70

```

```

QY      64  TNLVVAAGOSQSYFLDADADGAERHLEFGTSSSLPEFGSYNDLERVGH -RDOIPLG 122
      || ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
Db      71  TNAIVVAVRAGTOSTFLDADASSASDYLFTGTDDHSLPEFGYDRLERMAHQSPILG 130
QY      123  IEELIOSVSAALRYPGGSTRAOARSLILIQMISEARFNPIFMRAROYINSGESFLDMY 182
      :: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      131  LQALTHGISLFFRRSGNDMEEKARLIYLIVQVALEAARFRTYSNKNVRYSIQGTAFQDDA 190
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      183  MLEETSNKGQOSTOVQOSTDGVFNPNPRLGISTGNFVTLSNVRD-----VIASL 231
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      191  MISTLENNM-DMLRGVQESVDCTFPNQ-----VTLNFRREPIVYDLSLHPTAVVL 239
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      232  AIMLFLVC 238
      | : |||||
Db      240  AIMLFLVC 246

```

RESULT 6
 US-08-378-761A-77
 Sequence 77, Application US/08378761A
 Patent No. 5635384
 GENERAL INFORMATION:
 APPLICANT: WALSH, TERENCE A
 APPLICANT: HEY, TIMOTHY D
 APPLICANT: MORGAN, ALICE ER
 TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
 TITLE OF INVENTION: PRECHSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
 USING
 NUMBER OF SEQUENCES: 81
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ANDREA T. BORUCKI
 STREET: 9330 ZIONSVILLE ROAD
 CITY: INDIANAPOLIS
 STATE: IN
 COUNTRY: US
 ZIP: 46268
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/378, 761A
 FILING DATE: 26-JAN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: BORUCKI, ANDREA T
 REGISTRATION NUMBER: 33651
 REFERENCE/DOCKET NUMBER: 38272B
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (317) 337-4846
 INFORMATION FOR SEQ ID NO: 77:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 540 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-378-761A-77

[illegible]

Db	133	RENELTGPEDRIASLALYYITSCGIDPIPLANSFVNCIOMISEARFOIIEGEMTRR	192
Qy	173	SGSEFLPDWMLLETSWGQOSTOVQOSTDGVENNPFRLIGTGNFVLTSSNVDYATSLA	232
Db	193	YNNRSPADPSVITLSENGMRSLTAIOESNGAFASPIQLORRNGSKFNVDVSLIPITA	252
Qy	233	IMLFVCDRPSSDVRYWPLVIRPVLENSAVVDVYCTASPEVRYIV	279
Db	253	LMVRCAPPSQ---FSLIRPVVFNFNA--DV-CMDEPILVRI	292

RESULT 7

```

US-08-485-286-77
; Sequence 77, Application US/08485286
; Patent No. 5646026
; Patent No. 5646026 5646119
GENERAL INFORMATION:
APPLICANT: WALSH, TERENCE A
APPLICANT: HEY, TIMOTHY D
APPLICANT: MORAN, ALICE ER
TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
TITLE OF INVENTION: USING
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ANDREA T. BORUCKI
STREET: 9330 ZIONSVILLE ROAD
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: US
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,286
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/378761
FILING DATE: 26-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: BORUCKI, ANDREA T
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 38272B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 337-4846
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-286-77

```

	Query Match	29.7%;	Score 423;	DB 1;	Length 540;	
	Best Local Similarity	39.7%;	Pred. No. 1.7e-36;			
	Matches 114; Conservative	45;	Mismatches 106;	Indels 22;	Gaps 10;	
OY	4 TGDEFFRFTLLRLDYSSGS-FSNELPL-RSTIPVSDAQFVLVELTNOGDSITAI	61				
	: : : : : : : : : : : : : : : : : : :					
Dd	17 TWESTYNFRRAVRSHLTGCADVRHEIPIVLPNREGLPIS--ORFLIVELSNAHLSYTAL	74				
OY	62 DVTNLVVAIQAGDQSYFLRDAPDAE-----RHFTGTTRS-SLPFTSSTDLERYAGH	115				
	: : : : : : : : : : : : : : : : : : :					
Dd	75 DVTNAVYVCGCRAGNSAFEFH--PDNQDEAETHLFTFDVQNSFTFAFGNYNRDLQLGL	132				
OY	116 RDQIFLGIELIQSIQSALRY---PGGSTAAQRKSLIIILQMSEARFPPIWRARQYIN	172				
	: : : : : : : : : : : : : : : : : : :					

Db 133 REMIEIGTGPLEDALISALYYSTGCPQIPITLASFVNCIOMISEARFOYIEGEGKRIAIR 192

Qy 173 SGSEIIPDMYMLEETSMGQOSTOVQSTGVENNEPFRIGISGNFVTLNSVRDIYASLA 233

Db 193 YNRKSAPEDSVITLLENSMGRIETALQESNCGAASPIQLQRRNGKSNFYVDYSLIPIIA 252

Qy 233 IMLFVCRDRSSDVRNPELYIRPVLENSGAVDDYCTASPEYRIV 279

Db 253 LMYRCAPPPSSQ---FSLIRIVAPRNFNA--DY-CMDPEIRIV 292

RESULT 8

```

US-08-356-786-10
: Sequence 10, Application US/08356786
: Patent No. 5877305
: GENERAL INFORMATION:
: APPLICANT: Huston, James S.
: APPLICANT: Oppermann, Hermann
: APPLICANT: Houston, L. L.
: APPLICANT: Ring, David B.
: TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
: TITLE OF INVENTION: Marker
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
: STREET: Exchange Place, 53 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/356,786
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/831,967
: FILING DATE: 06-FEB-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Pitcher, Edmund R.
: REGISTRATION NUMBER: 27,829
: REFERENCE/DOCKET NUMBER: CAP-053
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 534 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-356-786-10

```

[illegible]

[illegible]

RESULT 9

Sequence 1 Application US/07901707
Patent No. 5376546
GENERAL INFORMATION:
APPLICANT: Bernhard, Susan L.
APPLICANT: Belter, Marc D.
APPLICANT: Carroll, Steve F.
APPLICANT: Lane, Julie A.
TITLE OF INVENTION: Materials Comprising and Methods of
Composition and Use for Ribosome-Inactivating Proteins
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/901,707
FILING DATE: 19920619
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5376546and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27129/30910
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-5750
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-901-707-1

Query Match	26.9%;	Score 383;	DB 1;	Length 267;
Best Local Similarity	39.9%;	Pred. No. 1e-32;		
Matches 101; Conservative	42;	Mismatches 94;	Indels 16;	Gaps 8;

QY	4	TGCEFFRITLLRDYSSGS-FSENEPLL-RQSTIPASDQRFVLELTMOGSDSTAAT	61
		: : : : : : : : : : : : : : : : : : : :	
Db	17	TVGSTNFIRAVRCRLTTGADVHRFVLEPNVNGRLPIN-QRRIVLELSHAELSTYAL	74
		: : : : : : : : : : : : : : : : : : : :	
QY	62	DYNTLVYVAOAGQSYELDADPQAE----RHLEFGT-TRESLPFGSYTLDERAGH	115
		: : : : : : : : : : : : : : : : : : : :	
Db	75	DYNTALYVVGTRAGNSATFEH-FDNOQDAALTHLEFDVONRTFAAGGNTDRLEQJAGN	132
		: : : : : : : : : : : : : : : : : : : :	
QY	116	RDQIPGLIEELLQSVALKR---PGGSTRQAQRSLIILQMTSEARFNPIFWARQYT	171
		: : : : : : : : : : : : : : : : : : : :	
Db	133	LRENIETLGNCPLEPAISALYVSTGGQLPTLARSFLICQMTSEARFQYIGCEAMKTR	192
		: : : : : : : : : : : : : : : : : : : :	

OY NGSGSPLEDDMTELEFMSGOOSQVOOSTDGVNNEPFRIGISNGVTSJNVADYAT 231
||| : : : : :
Db RYNRSAPDEPVITLHNSWGLSTAIQESNGGAFASFIQLRRNGSKFYVDYSILPII 252

OY 232 AIMEFVCRRDPS 244
||| : : : : :
Db 253 ALMYRCAPPPS 265

RESULT 10

```

US-07-988-430-1
: Sequence 1, Application US/07988430
: Patent No. 5416202
: GENERAL INFORMATION:
: APPLICANT: Bernhard, Susan L.
: APPLICANT: Better, Marc D.
: APPLICANT: Carroll, Stephen F.
: APPLICANT: Lane, Julie A.
: APPLICANT: Lei, Shau-Ping
: TITLE OF INVENTION: Materials Comprising and Methods of
: PREPARATION AND USE FOR RIBOSOME-INACTIVATING PROTEINS
: NUMBER OF SEQUENCES: 101
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: ADDRESSEE: Bicknell
: STREET: Two First National Plaza, 20 South Clark
: STREET: Street
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/988,430
: FILING DATE: 19921209
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/901,707
: FILING DATE: 19-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/787,567
: FILING DATE: 04-NOV-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: NO. 5416202and, Greta E.
: REGISTRATION NUMBER: 35302
: REFERENCE/DOCKET NUMBER: 31133
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 346-5750
: TELEFAX: (312) 984-9740
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 267 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-07-988-430-1

```

Query Match	26.98;	Score 383;	DB 1;	Length 267;
Best Local Similarity	39.98;	Pred. No. 1e-32;		
Matches 101;	Conservative 42;	Mismatches 94;	Indels 16;	Gaps 8;

[illegible]

Db 75 DVTNAYVVGVRAGNSAYFFH--PDNQEDAETHLFTDVQNKRYTFAGGNDYRLQOLAGN 132
QY 116 -RDQIPLGIEELIQSVSALRY---PGGSTRQAQARSLIILQIMISANPNPIEMRQYI 171
Db 133 LRENIEELGNPLEAIEAIALYYSTGTQOLPTLARSPFIQIMISANRFOYIEGEMRTRI 192
QY 172 NSGESFLPDMYMLEETSMGQOSTVOVQOSTGVFNNPRLGISTGNFTLSNVRVYIASL 231
Db 193 RYNRSAPDPSPVITLSENGRSLSTAIOESNOGAFASPIQOLQRNGSKFSYDVSTILPII 252
QY 232 AIMLFVCRDRPSS 244
Db 253 ALMYRCAPPSS 265

RESULT 11
US-08-218-303-16
; Sequence 16, Application US/08218303
; Patent No. 5547867
; GENERAL INFORMATION:
; APPLICANT: Kara, Bhupendra V.
; APPLICANT: Hockney, Robert C.
; APPLICANT: Filton, John E.
; TITLE OF INVENTION: FERMENTATION PROCESS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby & Cushman
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,303
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/841,533
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kokulis, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: PKR/3893/94908/MJW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO. 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-218-303-16

Query Match 26.9%; Score 383; DB 1; Length 267;
Best Local Similarity 39.9%; Pred. No. 1e-32;
Matches 101; Conservative 42; Mismatches 94; Indels 16; Gaps 8;

QY 4 TGDEYFRFTLLRDYVSSGS-FSNEIPL-LRSTIPVSDAQRVVLVELTNOGSDSTIAI 61
Db 17 TVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPIN--QRILVELSNHAELSVTLAL 74
QY 62 DVTNLYVAVQAGDOSYFLRDAPDGAE-----RHLEFTGT-TRSSLPTFGSTYDLERYAGH 115
Db 75 DVTNLYVAVQAGDQSYFLRDAPDGAE-----RHLEFTGT-TRSSLPTFGSTYDLERYAGH 132
QY 116 -RDQIPLGIEELIQSVSALRY---PGGSTRQAQARSLIILQIMISANPNPIEMRQYI 171

Db 133 LRENIEELGNPLEAIEAIALYYSTGTQOLPTLARSPFIQIMISEARFOYIEGEMRTRI 192
QY 172 NSGESFLPDMYMLEETSMGQOSTVOVQOSTGVFNNPRLGISTGNFTLSNVRVYIASL 231
Db 193 RYNRSAPDPSPVITLSENGRSLSTAIOESNOGAFASPIQOLQRNGSKFSYDVSTILPII 252
QY 232 AIMLFVCRDRPSS 244
Db 253 ALMYRCAPPSS 265

RESULT 12
US-08-425-336-1
; Sequence 1, Application US/08425336
; Patent No. 5621083
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroli, Stephen F.
; APPLICANT: Studulka, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Rbdsome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,336
; FILING DATE: 18-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 31394
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-425-336-1

Query Match 26.9%; Score 383; DB 1; Length 267;
Best Local Similarity 39.9%; Pred. No. 1e-32;
Matches 101; Conservative 42; Mismatches 94; Indels 16; Gaps 8;

QY 4 TGDEYFRFTLLRDYVSSGS-FSNEIPL-LRSTIPVSDAQRVVLVELTNOGSDSTIAI 61
Db 17 TVQSYTNFIRAVRGLTTGADVREIPLVLPNRVGLPIN--QRILVELSNHAELSVTLAL 74
QY 62 DVTNLYVAVQAGDOSYFLRDAPDGAE-----RHLEFTGT-TRSSLPTFGSTYDLERYAGH 115
Db 116 -RDQIPLGIEELIQSVSALRY---PGGSTRQAQARSLIILQIMISANPNPIEMRQYI 171

Db 75 DVTNAYVGYRAGNSAYFEH--PDNQEDAETHLFTDVQNRRTFAFGNYDRLQLAGN 132
 Qy 116 -RDQIPGIEELIOSVALRY---PGGSTRQAQARSLIILMIISARFNPITFMRAROYI 171
 Db 133 LRENIELNGPLLEAIALYYSSTGTQPLTARSFIIICMIISARFQYIEGEMRTRI 192
 Qy 172 NSGSEFLPDWYMLELETSMGQOSTGVFNNPRLGISTGNFTLSNVRDIASL 231
 Db 193 RYNRSAPDPVSITLSEMSGRISTAIQESNOGAFASPIQLQRRNGSKFSYDVSLIPII 252
 Qy 232 AIMEVCRDRPSS 244
 Db 253 ALMYRCAPPPSS 265

RESULT 13

US-08-488-113B-1
 ; Sequence 1, Application US/08488113B
 ; Patent No. 5744580

GENERAL INFORMATION:

APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studnika, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 169
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/488,113B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/425,336
 FILING DATE: 18-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/064,691
 FILING DATE: 12-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 267 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-488-113B-1

Query Match 26.98: Score 383, DB 1: Length 267;
 Best Local Similarity 39.98: Pred. No. 1e-32;
 Matches 101; Conservative 42; Mismatches 94; Indels 16; Gaps 8;

Qy 4 TGDEYFRPITLRLDYSSGS-FSNEIPL-LRQSTIPYSDAORFVLTNQGDSITAI 61
 Db 17 TVQSYTFIRAVRKRLTTGADVHEIPVLPNRGLPIN--QRIVELSHAEISVTAL 74
 Qy 62 DVTNAYVGYRAGNSAYFEH--PDNQEDAETHLFTDVQNRRTFAFGNYDRLQLAGN 132
 Db 75 DVTNAYVGYRAGNSAYFEH--PDNQEDAETHLFTDVQNRRTFAFGNYDRLQLAGN 132
 Qy 116 -RDQIPGIEELIOSVALRY---PGGSTRQAQARSLIILMIISARFNPITFMRAROYI 171
 Db 133 LRENIELNGPLLEAIALYYSSTGTQPLTARSFIIICMIISARFQYIEGEMRTRI 192
 Qy 172 NSGSEFLPDWYMLELETSMGQOSTGVFNNPRLGISTGNFTLSNVRDIASL 231
 Db 193 RYNRSAPDPVSITLSEMSGRISTAIQESNOGAFASPIQLQRRNGSKFSYDVSLIPII 252
 Qy 232 AIMEVCRDRPSS 244
 Db 253 ALMYRCAPPPSS 265

RESULT 14

US-08-477-484B-1
 ; Sequence 1, Application US/08477484B
 ; Patent No. 5756699

GENERAL INFORMATION:

APPLICANT: Better, Marc D.
 APPLICANT: Carroll, Stephen F.
 APPLICANT: Studnika, Gary M.
 TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 TITLE OF INVENTION: Proteins
 NUMBER OF SEQUENCES: 169
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,484B
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 03/425,336
 FILING DATE: 18-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 03/064,691
 FILING DATE: 12-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 PRIOR APPLICATION DATA:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 267 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-477-484B-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:25:55 ; Search time 15.4221 Seconds
(without alignments) 833.848 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1425
Sequence: 1 HQTGDEYFRFTLLRDYVS.....NSGAVDVTCTASPTVRIY 279

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCY_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCYUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1131	79.4	252	10	US-09-347-064-8
2	1127	79.1	252	10	US-09-347-064-2
3	328	23.0	247	9	US-09-792-793A-39
4	309	21.7	251	10	US-09-765-527-247
5	309	21.7	293	10	US-09-765-527-259
6	309	21.7	309	10	US-09-765-527-253
7	309	21.7	332	10	US-09-765-527-251
8	297	20.8	247	9	US-09-792-793A-34
9	249	17.5	250	9	US-09-792-793A-36
10	192	13.5	275	9	US-09-792-793A-35
11	182	12.8	254	9	US-09-792-793A-85
12	182	12.8	327	9	US-09-792-793A-85
13	182	12.8	330	9	US-09-792-793A-82
14	182	12.8	332	9	US-09-792-793A-82
15	182	12.8	332	9	US-09-792-793A-73
16	181	12.7	263	10	US-09-978-274A-4
17	181	12.7	314	10	US-09-978-274A-4
18	128.5	9.0	318	10	US-09-334-477-6
19	128.5	9.0	326	10	US-09-334-477-25

20	128	9.0	319	9	US-09-792-793A-38	Sequence 38, Appl
21	128	9.0	319	9	US-09-870-759-28	Sequence 28, Appl
22	128	9.0	694	10	US-09-334-477-49	Sequence 49, Appl
23	127	8.9	323	9	US-09-792-793A-80	Sequence 80, Appl
24	127	8.9	325	9	US-09-792-793A-81	Sequence 81, Appl
25	126.5	8.9	711	10	US-09-334-477-35	Sequence 35, Appl
26	125.5	8.8	329	10	US-09-334-477-39	Sequence 39, Appl
27	125	8.8	708	10	US-09-334-477-33	Sequence 33, Appl
28	124	8.7	326	10	US-09-334-477-37	Sequence 37, Appl
29	124	8.7	600	10	US-09-334-477-47	Sequence 47, Appl
30	122	8.6	325	9	US-09-792-793A-74	Sequence 74, Appl
31	122	8.6	327	9	US-09-792-793A-75	Sequence 75, Appl
32	121	8.5	247	9	US-09-792-793A-83	Sequence 83, Appl
33	121	8.5	249	9	US-09-792-793A-84	Sequence 84, Appl
34	121	8.5	320	9	US-09-792-793A-77	Sequence 77, Appl
35	121	8.5	322	9	US-09-792-793A-78	Sequence 78, Appl
36	121	8.5	325	9	US-09-792-793A-71	Sequence 71, Appl
37	121	8.5	327	9	US-09-792-793A-72	Sequence 72, Appl
38	120.5	8.5	293	9	US-09-792-793A-37	Sequence 37, Appl
39	120.5	8.5	315	10	US-09-334-477-2	Sequence 2, Appl
40	120.5	8.5	323	10	US-09-334-477-21	Sequence 21, Appl
41	116	8.1	110	10	US-09-978-274A-8	Sequence 8, Appl
42	90.5	6.4	400	10	US-09-895-211-4	Sequence 4, Appl
43	90.5	6.4	400	10	US-09-895-211-6	Sequence 6, Appl
44	86	6.0	1543	9	US-10-005-338B-8	Sequence 8, Appl
45	85	6.0	454	10	US-09-815-242-11157	Sequence 11157, A

ALIGNMENTS

```
RESULT 1
US-09-347-064-8
; Sequence 8, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; EARLIER FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-8
Query Match          79.4%; Score 1131; DB 10; Length 252;
Best Local Similarity 90.5%; Pred. No. 3e-104;
Matches 220; Conservative 10; Mismatches 13; Indels 0; Gaps 0;
QY 1 HQTGDEYFRFTLLRDYSSGSFNEIPLLRQSTIPVSDAORFVLVELTNOGGSTTAA 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 10 HQTGDEYFRFTLLRDYSSGSFNEIPLLRQSTIPVSDAORFVLVELTNOGGSTTAA 69
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 IVYTNLYVYVAYAGQSTFLRDAPDGAERHLLFTGTRSSLPFTGSYTLDERVAGHRDIP 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 70 IDVTMLYVYVAYAGQSTFLRDAPDGAERHLLFTGTRSSLPFTGSYTLDERVAGHRDIP 129
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 121 LGTEELIOSVSLARYPGGSTRACARSLITLTIOMISEARFPNIFPARAQYINSGESFLPD 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 130 LGIDDLIQSVATLRPPGGSTRTQARSILITLIOMISEARFPNIFPARAQYINSGASFLPD 189
```

QY 181 MYMLEETSMGQOSTQVQSTGCVFNNPRLGISTGNEVYTLNVDYVTLAIAIMLFVCD 240
 :|||||
 Db 190 VYMLEETSMGQOSTQVQSTGCVFNNPRLAIPGNEVYTLNVDYVTLAIAIMLFVCGE 249
 QY 241 RPS 243
 :|||
 Db 250 RPS 252

RESULT 2

US-09-347-064-2
 ; Sequence 2, Application US/09347064A
 ; Patent No. US20020045208A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eck, Jurgen
 ; APPLICANT: Schmidt, Arno
 ; APPLICANT: Zinke, Holger
 ; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
 ; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
 ; FILE REFERENCE: album
 ; FILE REFERENCE: 09282-5
 ; CURRENT APPLICATION NUMBER: US/09/347,064A
 ; EARLIER FILING DATE: 1999-07-02
 ; EARLIER APPLICATION NUMBER: PCT/EP98/00009
 ; EARLIER FILING DATE: 1998-01-02
 ; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 252
 ; TYPE: PRT
 ; ORGANISM: Viscum album
 ; US-09-347-064-2

Query Match 79.1%; Score 1127; DB 10; Length 252;
 Best Local Similarity 90.5%; Pred. No. 7.5e-104;
 Matches 219; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 HQTGDEYFRITLLRDYVSSGSFNSNIPILRQSTIPVSDAQRVYVELTNOGDSITAA 60
 :|||||
 Db 11 HQTGEEYFRITLLRDYVSSGSFNSNIPILRQSTIPVSDAQRVYVELTNOGDSITAA 70
 QY 61 IDVNVIVAAQADQSTFLDAPDGAERHLEFTGTTRSSLPFTGSYTDLEERYAGHRDQIP 120
 :|||||
 Db 71 IDVNLVYAAQADQSTFLDAPDGAERHLEFTGTTRSSLPFTGSYTDLEERYAGHRDQIP 130
 QY 121 LGIELIOSVSAALRYPGSGSTRAQARSLIILQIMISEARFNPIFWRARQYINSGESFLPD 180
 :|||||
 Db 131 LGIDQLIOSVTAALRYPGSGSTRAQARSLIILQIMISEARFNPIFWRARQYINSGESFLPD 190
 QY 181 MYMLEETSMGQOSTQVQSTGCVFNNPRLGISTGNEVYTLNVDYVTLAIAIMLFVCD 240
 :|||||
 Db 191 VYMLEETSMGQOSTQVQSTGCVFNNPRLAIPGNEVYTLNVDYVTLAIAIMLFVCGE 250
 QY 241 RP 242
 :||
 Db 251 RP 252

RESULT 3

US-09-792-793A-39
 ; Sequence 39, Application US/09792793A
 ; Patent No. US20020168370A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McDonald, John R.
 ; APPLICANT: Coggin, Philip
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
 ; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
 ; FILE REFERENCE: 25020-601D
 ; CURRENT APPLICATION NUMBER: US/09/792,793A
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 93

SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 39
 ; LENGTH: 247
 ; TYPE: PRT
 ; ORGANISM: Trichosantheus kirilowii
 ; US-09-792-793A-39

Query Match 23.0%; Score 328; DB 9; Length 247;
 Best Local Similarity 36.4%; Pred. No. 8.2e-25;
 Matches 87; Conservative 55; Mismatches 83; Indels 14; Gaps 9;

QY 4 TGDEYFRITLLRDYVSSGSFNSNIPILRQSTIPVSDAQRVYVELTNOGDSITAAIDV 63
 :|||||
 Db 10 TSSYGVFLSNLRKALPNERKLYDIPILR-SLPGS-QRYALHLTNVDETSTVAIDV 66
 QY 64 TNLVYAAQADQSTFLDAPDGAERHLEFTGTTRSSLPFTGSYTDLEERYAGHRDQIP 120
 :|||||
 Db 67 TNLVYIMGRAGDTSYFNEFASATEAKYVFDAMKVTLPYSGNERLQTAAGKIRENIP 126
 QY 121 LGIELIOSVSAALRYPGSGSTRAQARSLIILQIMISEARFNPIFWRARQYINSGESFLPD 180
 :|||||
 Db 127 LGELPLDSAITLTFYNNANSAASA-LMVLQSTSEARAKYFIQOIGKRYD-KTFLPS 182
 QY 181 MYMLEETSMGQOSTQVQSTGCVFNNPRLGISTGNEVYTLNVDYVTLAIAIMLFVCD 240
 :|||||
 Db 183 LATISLENSWALSQKQIQAISTNNGQFESPVLVINAQORVYTLNVDYVTSNAILLL 241

RESULT 4

US-09-765-527-247
 ; Sequence 247, Application US/09765527
 ; Patent No. US2002000638A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Better, Marc D.
 ; TITLE OF INVENTION: Methods for Recombinant Microbial Production of
 ; TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides

NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/765,527
 FILING DATE: 18-Jan-2001
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/621,803
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Borun, Michael F.
 REGISTRATION NUMBER: 25,447
 REFERENCE/DOCKET NUMBER: 27129/33199
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 247:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 251 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 247:

US-09-765-527-247
 Query Match 21.7%; Score 309; DB 10; Length 251;

Best Local Similarity 35.9%; Pred. No. 6.3e-23;
Matches 85; Conservative 37; Mismatches 107; Indels 8; Gaps 4;

QY 8 YFRFTLLRDYVSSGSGSNEIPLRQSTIPVSDAQRFVLVELTNOGDSITRAIDVTNLY 67
Db 17 YVNFLELRYKLKPEGNSHGIPLRKC--DDPGKCFVVALSNDNGQLAEIAIDVTSY 74
QY 68 VVAYQAGDOSYFLRDPDGAERHLFTGTTRSSLPFTGSDYDLERYAGHRDQIPLGIEEL- 126
Db 75 VVGVOYRNRSYFFKADAPDAIEGLFKNTIKTRLHFGSYPSLEGEKAYRETTDLGIEPLR 134
QY 127 --IQVSALRYPEGSTRAOARSLIILQIMSEARFNPFWRARQYINSGESLPLDMYL 184
Db 135 IGKKLDENAIIDNKPEIASLLVYQWSEARFTEINQIRN--NFOQRIKRPANNTI 192
QY 185 ELETSMGQOSTVOQOS--TDGVFNPNPRLGISTGNFVTLNVRDYASLAIMLFVCHDRPS 240
Db 193 SLEKMKGLSFQIRTSANGMFSSEAVLELRANCKRYVAVDQVKRIALLKFFVDKD 249

RESULT 5
US-09-765-527-259

; Sequence 259, Application US/09765527
; Patent No. US20020006638A1
; GENERAL INFORMATION:

APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803

FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 259:
SEQUENCE CHARACTERISTICS:
LENGTH: 293 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 259:

US-09-765-527-259

Query Match 21.7%; Score 309; DB 10; Length 293;
Best Local Similarity 35.1%; Pred. No. 7.8e-23;
Matches 85; Conservative 40; Mismatches 109; Indels 8; Gaps 4;

QY 8 YFRFTLLRDYVSSGSGSNEIPLRQSTIPVSDAQRFVLVELTNOGDSITRAIDVTNLY 67
Db 39 YVNFLELRYKLKPEGNSHGIPLRKC--DDPGKCFVVALSNDNGQLAEIAIDVTSY 96
QY 68 VVAYQAGDOSYFLRDPDGAERHLFTGTTRSSLPFTGSDYDLERYAGHRDQIPLGIEEL- 126

Db 97 VVGVOYRNRSYFFKADAPDAIEGLFKNTIKTRLHFGSTYPSLEGEKAYRETTDLGIEPLR 156

QY 127 --IQVSALRYPEGSTRAOARSLIILQIMSEARFNPFWRARQYINSGESLPLDMYL 184
Db 157 IGKKLDENAIIDNKPEIASLLVYQWSEARFTEINQIRN--NFOQRIKRPANNTI 214

QY 185 ELETSMGQOSTVOQOS--TDGVFNPNPRLGISTGNFVTLNVRDYASLAIMLFVCHDRPS 243
Db 215 SLEKMKGLSFQIRTSANGMFSSEAVLELRANCKRYVAVDQVKRIALLKFFVDKDRS 274

QY 244 SS 245
Db 275 AA 276

RESULT 6
US-09-765-527-253

; Sequence 253, Application US/09765527
; Patent No. US20020006638A1
; GENERAL INFORMATION:

APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803

FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 253:
SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 253:

US-09-765-527-253

Query Match 21.7%; Score 309; DB 10; Length 309;
Best Local Similarity 35.1%; Pred. No. 8.4e-23;
Matches 85; Conservative 40; Mismatches 109; Indels 8; Gaps 4;

QY 8 YFRFTLLRDYVSSGSGSNEIPLRQSTIPVSDAQRFVLVELTNOGDSITRAIDVTNLY 67
Db 39 YVNFLELRYKLKPEGNSHGIPLRKC--DDPGKCFVVALSNDNGQLAEIAIDVTSY 96
QY 68 VVAYQAGDOSYFLRDPDGAERHLFTGTTRSSLPFTGSDYDLERYAGHRDQIPLGIEEL- 126
Db 97 VVGVOYRNRSYFFKADAPDAIEGLFKNTIKTRLHFGSTYPSLEGEKAYRETTDLGIEPLR 156
QY 127 --IQVSALRYPEGSTRAOARSLIILQIMSEARFNPFWRARQYINSGESLPLDMYL 184

us-09-627-165e-16.rapb

```

QY 8 XFFATLLRROYSSGSFSEIPLLRQSTIPVSDAQREVELVETNOGDSITPAIDVYTLX 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18 YLSTTIRKRYKVAADKTEOCTI-----OKISKFTGRSYIDLVSSTQKTIATIMADLY 72
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 68 VVAY----QAGDQSYFLRPAAPGARHFL--TGTSSSLPTFGSTYDLERYAGHRDIP 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 73 VLGYSDIANNKGAFEPFKRYVTEAVANNPFGATNTNRIKLTFTGSGYGLDEKNGLRKNP 133
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 LGEIELQSYSAIARPGGSTRQAARSLITLLOMISEARPNPIEMRQYINSG-----174/
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133 LGEIPLRENSIVINYGKAGVYKQAKFLALAIQWSEAR-----KITSDKIPSEKY 184
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 175 ESFLPDMYMLETSSMGQSOYQOOS 200
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 185 EEVTVDEYMTALENNMAKLSTAVYNS 210
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Query Match	13.58	Score 192	DB 9	Length 275
Best Local Similarity	26.2%	Pred. No. 2,7e-11		
Matches 75	Conservative 56	Mismatches 99	Indels 56	Gaps 15
QY	4	TGDEFREFTLLRDYVSSGSFSNEITLQ-----STPVSDAQDFVLLVETLNOGDSIT	58	
Db	12	TAGGVSSVVKIRINNVAD-----PMLYGGGTIDIAVGPPSKKEFLRINQSSRG-TVS	63	
QY	59	AAIDVTMLYVAVQAGD-----QSYFLRDAPOGAE-RHLF---TGTRRSILPTGTSTDI	109	
Db	64	LGLRDNMLYVAVYALMDNTVNRVARYFRREITISASSTLLEPAAVNAOKALEVTEYDQSI	123	
QY	110	ERYA-----GHRDQPLPIGIEELISVSALRPGGSTRQAQRSLITLQIOMSEAKRFNPI	163	
Db	124	EKNMOITQGDOSKRELGIDILSTSMEXVANKKARVYKDEARFLILIAQMTAFANAF---	180	
QY	164	FWARQRY-----NGESFLLPDMTMLELETSWGQOSTOYO-OSTDGVEFNNPRLGISTG	216	
Db	181	-----RIQNLVIKNNPPKNFNSKNVVIDFEVNMKKISTAIYGAOKNGVEFNKDYDEG----	231	
QY	217	NEVTLSVNRDVATSLAIMLFWCRDRPS---SSDVR-TWPLVIRPVYL	258	
Db	232	-FGKVRQVQDL-QMGILMYLGKPRSSSEANSTYRAHGCVL--KPTL	272	

```

? CURRENT FILING DATE: 2001-02-22
? NUMBER OF SEQ. ID NOS: 93
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 85
? LENGTH: 254
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Methionine-Saporin fusion
US-09-792-793A-85

```

```

? RESULT 12
? US-09-792-793A-79
? Sequence 79, Application US/09/792793A
? Patent No. US20020168370A1
? GENERAL INFORMATION:
? APPLICANT: McDonald, John R.
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
? FILE REFERENCE: 25020-601D
? CURRENT APPLICATION NUMBER: US/09/792,793A
? CURRENT FILING DATE: 2001-02-22
? NUMBER OF SEQ ID NOS: 93
? SOFTWARE: Patentln Ver. 2.0
? SEQ ID NO 79
? LENGTH: 327
? TYPE: PR1
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Chemokine-toxin fusion protein
? US-09-792-793A-79

```

	Query Match	12.8%	Score 192;	DB 9;	Length 327;
	Best Local Similarity	26.3%	Ped. NO.	3.3e-10;	
	Matches	68;	Conservative	47;	Mismatches 96; Indels 48; Gaps 11;
OY	4	TGDEFFRITLLDNYSSGSFSLPLRLQ-----STPVSDAQRFVLVLTNOGDSIT	58		
	: :	: :	:	:	:
Dd	86	TAGGYSSFEVKIKNNKYD-----PNKYGQTDAIVGPSPKEKFLRNFGSRR-TYS	137		
OY	59	AALDVNTLVVAYOAGD-----QSFEILDAPDGA-E-RHLF--TGTRSSLPTSGSYDL	109		
	: :	:	:	:	:
Dd	138	LGLKRDVLVVAILANDNNTNVRAITYESETTSAESLTPLEATTANOKALETETYDCSI	197		
OY	110	ERVA-----GHRDQIPIGIELLIOSVALRYPGGSTPRACRSILILLIOMISEARFNPI	163		
	: :	:	:	:	:

Thu Mar 27 07:07:01 2003

us-09-627-165e-16.rapb

Page 6

Db 198 EKNAQITGDSRKELGIDILSTSMKAVNKKARVDEARFLLIAIOMTAEARF--- 254
Qy 164 FWRARQYI-----NSGESFLPDMYMLETSMGOOSTOVO-OSTDGVFNNPRLGISTG 216
Db 255 -----RYIONLVKKNFPNKNFNSKNVIOFEVNMKKISTAIYGDANKGVFNKDYDFG---- 305
Qy 217 NFVTLISNVRDVIASLAIML 235
Db 306 -FGKRVQVKDLOMGLMLMYL 323

RESULT 13
US-09-792-793A-82
; Sequence 82, Application US/09/792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 82
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chemokine-toxin fusion prote
US-09-792-793A-82

Query Match 12.8%; Score 182; DB 9; Length 330;
Best Local Similarity 26.3%; Pred. No. 3.4e-10;
Matches 68; Conservative 47; Mismatches 96; Indels 48; Gaps 11;

Qy 4 TGDYFRFTLLRDYVSSGFSNEIPILRQ-----STIPVSDARFVIVELTNOGDSIT 58
Db 89 TAGOYSSFDKIRNNVKD-----PNLKYGTDIAYIGPPSKKEFLRINFQSSRG-TVS 140
Qy 59 AALDVNTLVVAVYAGD-----QSYFLDAPDGAE-RHLF---TGTRSSLPTGSGYTDL 109
Db 141 LGLKRDNLVYVAYLAMDNTNVRAYFRSEITSAESTALPEATTANOKALEYTEDYOSI 200
Qy 110 ERYA-----GHRDQIPGLIEFLIOSVSLRYPGSGTRAQARSLLIILQIMISEARFNPI 163
Db 201 EKNAQITGDSRKELGIDILSTSMKAVNKKARVDEARFLLIAIOMTAEARF--- 257
Qy 164 FWRARQYI-----NSGESFLPDMYMLETSMGOOSTOVO-OSTDGVFNNPRLGISTG 216
Db 258 -----RYIONLVKKNFPNKNFNSKNVIOFEVNMKKISTAIYGDANKGVFNKDYDFG---- 308
Qy 217 NFVTLISNVRDVIASLAIML 235
Db 309 -FGKRVQVKDLOMGLMLMYL 326

RESULT 14
US-09-792-793A-73
; Sequence 73, Application US/09/792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 73
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chemokine-toxin fusion pro
US-09-792-793A-73

Query Match 12.8%; Score 182; DB 9; Length 332;
Best Local Similarity 26.3%; Pred. No. 3.4e-10;
Matches 68; Conservative 47; Mismatches 96; Indels 48; Gaps 11;

Qy 4 TGDYFRFTLLRDYVSSGFSNEIPILRQ-----STIPVSDARFVIVELTNOGDSIT 58
Db 91 TAGOYSSFDKIRNNVKD-----PNLKYGTDIAYIGPPSKKEFLRINFQSSRG-TVS 142
Qy 59 AALDVNTLVVAVYAGD-----QSYFLDAPDGAE-RHLF---TGTRSSLPTGSGYTDL 109
Db 143 LGLKRDNLVYVAYLAMDNTNVRAYFRSEITSAESTALPEATTANOKALEYTEDYOSI 202
Qy 110 ERYA-----GHRDQIPGLIEFLIOSVSLRYPGSGTRAQARSLLIILQIMISEARFNPI 163
Db 203 EKNAQITGDSRKELGIDILSTSMKAVNKKARVDEARFLLIAIOMTAEARF--- 259
Qy 164 FWRARQYI-----NSGESFLPDMYMLETSMGOOSTOVO-OSTDGVFNNPRLGISTG 216
Db 260 -----RYIONLVKKNFPNKNFNSKNVIOFEVNMKKISTAIYGDANKGVFNKDYDFG---- 310
Qy 217 NFVTLISNVRDVIASLAIML 235
Db 311 -FGKRVQVKDLOMGLMLMYL 328

RESULT 15
US-09-792-793A-76
; Sequence 76, Application US/09/792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE
; TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
; FILE REFERENCE: 25020-601D
; CURRENT APPLICATION NUMBER: US/09/792,793A
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 76
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chemokine-toxin fusion pr
US-09-792-793A-76

Query Match 12.8%; Score 182; DB 9; Length 332;
Best Local Similarity 26.3%; Pred. No. 3.4e-10;
Matches 68; Conservative 47; Mismatches 96; Indels 48; Gaps 11;

Qy 4 TGDYFRFTLLRDYVSSGFSNEIPILRQ-----STIPVSDARFVIVELTNOGDSIT 58
Db 91 TAGOYSSFDKIRNNVKD-----PNLKYGTDIAYIGPPSKKEFLRINFQSSRG-TVS 142
Qy 59 AALDVNTLVVAVYAGD-----QSYFLDAPDGAE-RHLF---TGTRSSLPTGSGYTDL 109
Db 143 LGLKRDNLVYVAYLAMDNTNVRAYFRSEITSAESTALPEATTANOKALEYTEDYOSI 202
Qy 110 ERYA-----GHRDQIPGLIEFLIOSVSLRYPGSGTRAQARSLLIILQIMISEARFNPI 163
Db 203 EKNAQITGDSRKELGIDILSTSMKAVNKKARVDEARFLLIAIOMTAEARF--- 259

QY 164 FWRAROYI-----NSGESEFLPDWYMLEFTSMGOOSTOYQ-QSTDCVFNPNPRIGISTG 216
Db 260 -----RYIONLVIKNFNPKENSENKVYIQFEVWKKISTAIYGDANKGVFNKDYDFG----- 310
QY 217 NEVTLSNVRDYIASLAIML 235
Db 311 -FGKVRQVKDLDQMGILMYL 328

Search completed: March 18, 2003, 08:28:38
Job time : 16.4221 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:23:55 ; Search time 23.1332 Seconds
(without alignments)
1159.439 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1425
Sequence: 1 HQTGDEYFRFTLLRDYVS.....NSGAVDVTCTASEPTVRLV 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1119	78.5	254	2 PD0018	mistletoe lectin I
2	514	36.1	528	1 TZLSA	abrin-a precursor
3	481	33.8	528	2 S32431	abrin-d precursor
4	481	33.8	562	2 S16022	abrin-c precursor
5	471.5	33.1	527	2 S32430	abrin-b precursor
6	458	32.1	251	2 C39761	abrin (clone 7.2)
7	427	30.0	570	2 S62627	agglutinin I precu
8	425.5	29.9	576	1 RLCSG	ricin D precursor
9	423	29.7	564	1 RLCSAG	agglutinin precursor
10	340.5	23.9	286	2 S25560	RNA N-glycosidase
11	336	23.6	289	1 RLUTZT	RNA N-glycosidase
12	333.5	23.4	286	2 JC4235	RNA N-glycosidase
13	331.5	23.3	289	2 JC5606	karasurin C - Tric
14	329	23.1	247	2 JC5032	karasurin-B - Tric
15	327	22.9	247	2 JU0393	karasurin - Mongol
16	324.5	22.8	245	2 JC4840	RNA N-glycosidase
17	311	21.8	316	2 JT0753	RNA N-glycosidase
18	296.5	20.8	286	1 RLPUGG	RNA N-glycosidase
19	291.5	20.5	278	2 S23519	beta-luffin - smoo
20	286.5	20.1	277	2 S22494	luffin N-glycosidase
21	280	19.6	250	2 JN0108	luffin-b - smooth
22	249	17.5	278	2 A39817	RNA N-glycosidase
23	203	14.2	313	2 S17757	RNA N-glycosidase
24	202.5	14.2	261	2 JE0401	antifurcal protein
25	186	13.8	294	2 S28439	RNA N-glycosidase
26	186	13.2	253	2 S28539	RNA N-glycosidase
27	187	13.1	283	2 S05205	RNA N-glycosidase
28	186	13.1	253	2 S28542	RNA N-glycosidase
29	184	12.9	272	2 JC4811	betavulgin - beet

30	182	12.8	253	2 S29331	RNA N-glycosidase
31	182	12.8	292	1 RLQHC2	RNA N-glycosidase
32	180.5	12.7	106	2 B39761	abrin (clone 3.7)
33	179	12.6	253	2 S28541	RNA N-glycosidase
34	161	11.3	253	2 A58923	RNA N-glycosidase
35	157.5	11.1	310	2 S46239	ribosome-inactivat
36	148.5	10.4	293	2 S17519	RNA N-glycosidase
37	143.5	10.1	236	2 S17932	RNA N-glycosidase
38	134	9.4	289	2 T12573	RNA N-glycosidase
39	129	9.1	319	2 S58343	Shiga-like toxin I
40	128.5	9.0	318	2 S01032	Shiga-like toxin I
41	128	9.0	319	2 S21940	hypothetical prote
42	128	9.0	319	2 E90779	Shiga toxin 2 subu
43	128	9.0	319	2 G85640	Shiga toxin 2 subu
44	127.5	8.9	320	2 S42608	Shiga-like toxin -
45	127	8.9	319	2 I76713	variant shiga-like

ALIGNMENTS

RESULT 1

PD0018
mistletoe lectin I A chain - Viscum album (fragment)

C:Species: Viscum album

C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 26-Aug-1999

C:Accession: PD0018

R:Eschenburg, S.; Krauspenhaar, R.; Mikhailov, A.; Stoeva, S.; Betzel, C.; Voelter, A.

Biochem. Biophys. Res. Commun. 247, 367-372, 1998

A:Title: Primary structure and molecular modeling of mistletoe lectin I from Viscum

A:Reference number: PD0018; MUID:98308123; PMID:9642133

A:Accession: PD0018

A:Molecule type: protein

A:Residues: 1-254 <ESC>

C:Superfamily: ricin; RNA N-glycosidase homology

F:7-246/Domain: RNA N-glycosidase homology <RMC>

Query Match 78.5%; Score 1119; DB 2; Length 254;

Beet Local Similarity 89.0%; Pred. No. 1.6e-93;

Matches 218; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY	1	HQTGDEYFRFTLLRDYVS	SFSFNEIPLRQSTIPVSDAQRFLVLTNQGDSITAA	60
DB	10	HQTGDEYFRFTLLRDYVS	SFSFNEIPLRQSTIPVSDAQRFLVLTNQGDSITAA	69
QY	61	IDVTMLVYVAYAGQSTFL	RDAPGARHLFTGTRSSLPPTGSYDLERYAGHRQIP	120
DB	70	IDVTNAVYVAYAGQSTFL	RDAPGARHLFTGTRSSLPPTGSYDLERYAGHRQIP	129
QY	121	IGIEELIOSVSLARYPGS	TRQAARSLITLQIMSEARFNPITPMARQYINSGESFLPD	180
DB	130	IGIDDIQIOSVSLARYPG	SSTRQAARSLITLQIMSEARFNPITPMARQYINSGESFLPD	189
QY	181	MYMLELTSWGOOSTQVO	STDGVFNPFRLGISTGNFVLSNNVDYIASLAIMFLVCRD	240
DB	190	VMLELTSWGOOSTQVO	STDGVFNPFRLGISTGNFVLSNNVDYIASLAIMFLVCRD	249
QY	241	RRSSS	245	
DB	250	RRSSS	254	

RESULT 2

TZLSA

abrin-a precursor - Indian licorice (fragment)

N:Contains: RNA N-glycosidase (EC 3.2.2.22)

C:Species: Abrus precatorius (Indian licorice)

C:Date: 31-Dec-1993 #sequence_revision 01-Aug-1997 #text_change 16-Jul-1999

C:Accession: S32429; JT0202; A39761; JC1398; S14472; S24133; S74110; S74111

R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.

J. Mol. Biol. 229, 263-267, 1993

A:Title: Primary structure of three distinct isoabirins determined by cDNA sequencing.

A:Reference number: S32429; MUID:9132798; PMID:8421313

A:Accession: S32429
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'E', 2-528 <HUN>
A:Cross-references: GB:M8344; NID:q166294; PIDN:AAA3624.1; PID:q166295
A:Note: the coding region for the sequence shown is preceded by an ATG codon
A:Note: residues 1-8 were derived from the synthesized primer
R:Funatsu, G.; Taguchi, Y.; Kamemoto, M.; Yanaka, M.
Agric. Biol. Chem. 52, 1095-1097, 1988
A>Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from
A:Reference number: J70202
A:Accession: J70202
A:Molecule type: protein
A:Residues: 1-201,203-251 <FUN>
A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
R:Evensen, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A>Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A:Reference number: A39761; MUID:q1201329; PMID:2016300
A:Accession: A39761
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 'E', 2-251 <EVE>
A:Cross-references: GB:X54872
A:Note: residues 1-8 were derived from the synthesized primer
R:Kimura, M.; Sumizawa, T.; Funatsu, G.
Biosci. Biotechnol. Biochem. 57, 166-169, 1993
A>Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic
A:Reference number: JCI398; MUID:q9169023; PMID:7763422
A:Contents: seeds
A:Accession: JCI398
A:Molecule type: protein
A:Residues: 261-347, 'T', 349-351, 'A', 353-357, 'L', 359-528 <KIM>
A:Experimental source: seed
R:Evensen, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL Data Library, October 1990
A:Description: Direct molecular cloning of two distinct abrin A-chains.
A:Reference number: S14471
A:Accession: S14472
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 'M', 2-251 <EV2>
A:Cross-references: EMBL:X54873; NID:q16090; PIDN:CA38655.1; PID:q16091
R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.
FEBS Lett. 309, 115-118, 1992
A>Title: The complete primary structure of abrin-a B chain.
A:Reference number: S24133; MUID:q2371656; PMID:1505674
A:Accession: S24133
A:Molecule type: protein
A:Residues: 262-297, 'Y', 299-426, 'L', 428-466, 'P', 468-482, 'L', 484-528 <CHE>
R:Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.
Eur. J. Biochem. 240, 564-569, 1996
A>Title: Probing the domain structure of abrin-a by tryptic digestion.
A:Reference number: S74110; MUID:q7008945; PMID:8856055
A:Accession: S74110
A:Molecule type: protein
A:Residues: 89-108,154-172 <LIN>
A:Experimental source: seed
A:Accession: S74111
A:Molecule type: protein
A:Residues: 262-276, 'X', 278-280,329-348;369-388;399-418 <LIW>
A:Experimental source: seed
A:Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which inhibits
taining receptors on the cell surface. The A and B chains are linked by a single disulfide
C:Keywords: ricin; rRNA N-glycosidase homology
C:Superfamily: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyrogutamic acid
F:1-251/Product: abrin-a chain A #status experimental <ACH>
F:7-246/Domain: rRNA N-glycosidase homology <RNG>
F:283-325/Product: abrin-a chain B #status experimental <RNG>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status repeats
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted

F:247-269,286-305,329-346,417-430,456-473/Dissulfide bonds: #status predicted
F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F:361,401/Binding site: carbohydrazate (Asn) (covalent) #status experimental
F:500,521/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 36.1%; Score 514; DB 1; Length 528;
Best Local Similarity 42.5%; Pred. No. 2,1e-38;
Matches 122; Conservative 45; Mismatches 88; Indels 32; Gaps 8;

OY 4 TGEDEFRTITLRLDYVSSGSFSNEPILRQSTIPYSDAQRVLYELTNOGSDSTAIDV 63
 |:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 13 TSQSKEPIEALREXL-RGGLIHDIPLVDPT-TLGERNRIITLSIDTESIEVGDI 70

OY 64 TNLVVAQAQGDSQSYFLRDAPDGAERHLFTGTTRSSLPFTSSTYDLERYAGH-RDOJPLG 122
 |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 71 TNAYVAARAGQSFLELDAPSASADYLEFTCTDDHSLEPFYTYGDLEWMAHQSRQIPLG 130

OY 123 IEELIQSVSLARYPGCGSTRQAQRSLIILOMISEARPNPFWRARQYINGESFLPMY 182
 ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 131 LQALTHGISIFFSGGNDMEERAKRLDIVIOVAEAARFPFRYSINRWRSIQGTAFQPPAA 190

OY 183 MEETSMGOGOSTVOQOSTDGVEFNPFRLIGTGFWLVSWRD-----VIASL 231
 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 191 MSTLENMNDNISRGVQESVDTFPNQ-----VLTIIRNRPVYDSLHPYAVAL 240

OY 232 AIMEFVCRRDPSSSDRYWPLVIREPVLENLSGAVDVDTCTASEPTVRI 278
 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 241 ALMLEFVC-NPPVANQS---PLLIRFIVEKS----KICSSRYEPVRI 279

RESULT 3
532431
abrin-d precursor - Indian licorice (fragment)
N:Contains: RNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C>Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 01-Aug-1997
C:Accession: S32431; S34408
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J.Mol. Biol. 229, 261-267, 1993
A>Title: Primary structure of three distinct isoabrins determined by cDNA sequencing.
A:Reference number: S32429; MUID:93133798; PMID:8421313
A:Accession: S32431

A:Molecule type: mRNA
A:Residues: 1-528 <HDN>
A:Cross-references: GB:M98346
R:Hung, C.; Lee, M.; Lee, T.; Lin, J.
submitted to the EMBL Data Library, March 1993
A:Reference number: S34408
A:Accession: S34408

A:Molecule type: mRNA
A:Residues: 1-169, C', 171-320, 'U', 322-528 <HUZ>
A:Cross-references: GB:M98346
C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating
The A and B chains are linked by a single disulfide bond, which is essential for toxin
F:1/Modified site: rRNA N-glycosidase homology
C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; protein; rRNA
F:1-351/Product: abrin-d chain A #status predicted <ACH>
F:7-246/Domin: rRNA N-glycosidase homology <RNG>
F:261-528/Product: abrin-d chain B #status predicted <BCH>
F:283-323,326-366,369-407,414-449,453-492,495-528/Region: 40-residue repeats
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F:74,113,199,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted
F:200,253,361,401,402/Binding site: carbohydrazate (Asn) (covalent) #status predicted
F:247-269,286-305,329-346,417-430,456-473/Dissulfide bonds: #status predicted
F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F:500,521/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted

Query Match 33.8%; Score 481; DB 2; Length 528;
Best Local Similarity 43.7%; Pred. No. 2,1e-35;
Matches 121; Conservative 41; Mismatches 103; Indels 12; Gaps 7;

4 TGDEYFRPFTLLRDYVSSGSFSNEPILRQSTIPYSDAQRFVLYELTNOGSDSTAIDV 63


```
Db 13 TSQSYKQFTEALRQRL-TGGLIHDPVLPDPF-TVEERNKRIYVELNSKESELEVGIDV 70
QY 64 TNLVVAAYQAGDSYFLRDAPDGAERHLFTGTRSSLPFTGSYDLERYAGH-RDQIPLG 122
Db 71 TNAVVAAYRAGOSYFLRDAPASASTYLFPTQRYSLRFDGSDYDLERMAHQTRREISLG 130
QY 123 IEELIOSVALRYPGSGSRQAARSLIILOMISEAARFNPFMRARQYINSGESFLPDMY 182
Db 131 LQALTHAISFLRSGASNDDEKARLVIITQMASEAARFYISNRGVSIKRTGAFQPDPA 190
QY 183 MLELETSWGQOSTQVOQSTDGVE--NNPFRLGISTGNEFTLSNVADVIASLAIMLFCRDR 241
Db 191 MLELENNMNDNLSCGVQOSYODTFEPNNVILSSINQPVVDSLSHPYAVIALMLFVC-NP 249
QY 242 PSSSDVRYMPLVIRPVLENSGAVDDVCTASPTVRI 278
Db 250 PNAQOS---PLLRISIVEES---KICSSRYEPTVRI 279
```

RESULT 4

```
S16022
abrin-c precursor - Indian licorice
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S16022
R:Wood, K.A.; Lord, J.M.; Wawrzynczak, E.J.; Platak, M.
Eur. J. Biochem. 198, 723-732, 1991
A:Title: Preproabrin: genomic cloning, characterisation and the expression of the A-chain
A:Reference number: S16022; MUID:91266957; PMID:2050149
A:Accession: S16022
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-562 <WOO>
A:Cross-references: EMBL:X55667; NID:g16084; PIDN:CAA39202.1; PID:g16085
C:Comment: Abirin consists of an A chain, which inhibits protein synthesis by inactivating
the A and B chains are linked by a single disulfide bond, which is essential for toxicity
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic acid
F:35-285/Product: abrin-c chain A #status predicted <ACH>
F:41-280/Domain: rRNA N-glycosidase homology <RNG>
F:295-562/Product: abrin-c chain B #status predicted <BGH>
F:317-359,360-400,403-441,448-483,487-526,529-562/Region: 40-residue repeats
F:35/Modified site: pyroglutamic carboxylic acid (Gln) (in mature form) #status predicted
F:108,147,229,230/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:198,201/Active site: Glu, Arg #status predicted
F:234,287,395,435,436/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:281,303,320-338,383-380,451-464,490-507/Disulfide bonds: #status predicted
F:322,346/Binding site: N-acetylglactosamine (Asp, Asn) #status predicted
F:334,555/Binding site: N-acetylglactosamine (Asp, Asn) #status predicted
```

Query Match 33.8% Score 481; DB 2; Length 562;

Best Local Similarity 43.7%; Pred. No. 2,2e-35; Matches 121; Conservative 41; Mismatches 103; Indels 12; Gaps 7;

```
QY 4 TGEYFRFTLLRDVYSGSFSNEIPLRQSTIPVSDAQRFLVELTNOGDSITPAIDV 63
Db 47 TSQSYKQFTEALRQRL-TGGLIHDPVLPDPF-TVEERNKRIYVELNSKESELEVGIDV 104
QY 64 TNLVVAAYQAGDSYFLRDAPDGAERHLFTGTRSSLPFTGSYDLERYAGH-RDQIPLG 122
Db 105 TNAVVAAYRAGOSYFLRDAPASASTYLFPTQRYSLRFDGSDYDLERMAHQTRREISLG 164
QY 123 IEELIOSVALRYPGSGSTRQAARSLIILOMISEAARFNPFMRARQYINSGESFLPDMY 182
Db 165 LQALTHAISFLRSGASNDDEKARLVIITQMASEAARFYISNRGVSIKRTGAFQPDPA 224
QY 183 MLELETSWGQOSTQVOQSTDGVE--NNPFRLGISTGNEFTLSNVADVIASLAIMLFCRDR 241
Db 225 MLELENNMNDNLSCGVQOSYODTFEPNNVILSSINQPVVDSLSHPYAVIALMLFVC-NP 283
QY 242 PSSSDVRYMPLVIRPVLENSGAVDDVCTASPTVRI 278
```

```
Db 284 PNAQOS---PLLRISIVEES---KICSSRYEPTVRI 313
```

RESULT 5

```
S32430
abrin-b precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 20-Aug-1999
C:Accession: S32430; JCI399
R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
J. Mol. Biol. 229, 263-267, 1993
A:Title: Primary structure of three distinct isoabirins determined by cDNA sequencing.
A:Reference number: S32429; MUID:93132798; PMID:8421313
A:Accession: S32430
A:Molecule type: mRNA
A:Residues: 1-527 <HUN>
A:Cross-references: GB:M98345; NID:g166296; PIDN:AAA32625.1; PID:g166297
R:Kimura, M.; Sumizawa, T.; Funatsu, G.
Biosci. Biotechnol. Biochem. 57, 166-169, 1993
A:Title: The complete amino acid sequences of the B-chains of Abirin-a and Abirin-b, to
A:Reference number: JCI398; MUID:93169023; PMID:7763422
A:Accession: JCI399
```

A:Molecule type: protein
A:Residues: 260-281, 'D', 283-290, 'N', 292-349, 'PQ', 352-377, 'N', 379-425, 'W', 427, 'D', 429-
A:Experimental status: seed
C:Superfamily: ricin; rRNA N-glycosidase homology
F:1-250/Product: abrin-b chain A #status predicted <ACH>
F:7-245/Domain: rRNA N-glycosidase homology <RNG>
F:260-527/Product: abrin-b chain B #status experimental <BGH>
F:282-324,325-365,368-406,413-448,452-491,494-527/Region: 40-residue repeats
F:1/Modified site: pyroglutamic carboxylic acid (Gln) #status predicted
F:110,113,119,194,195/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:110,360,400/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:163,166/Active site: Glu, Arg #status predicted
F:246,268,285-304,328-345,416-429,455-472/Disulfide bonds: #status predicted
F:287,311/Binding site: N-acetylglactosamine (Asp, Asn) #status predicted
F:499,520/Binding site: N-acetylglactosamine (Asp, Asn) #status predicted

Query Match 33.1% Score 471.5; DB 2; Length 527;

Best Local Similarity 42.6%; Pred. No. 1.5e-34; Matches 118; Conservative 46; Mismatches 100; Indels 13; Gaps 8;

```
QY 4 TGEYFRFTLLRDVYSGSFSNEIPLRQSTIPVSDAQRFLVELTNOGDSITPAIDV 63
Db 13 TSQSYKQFTEALRQRL-TGGLIHDPVLPDPF-TQERNRRIYVELNSDSTESIEAGIDV 70
QY 64 TNLVVAAYQAGDSYFLRDAPDGAERHLFTGTRSSLPFTGSYDLERYAGH-RDQIPLG 122
Db 71 TNAVVAAYRAGRSYFLRDAPTSASRYLFTGQOYSLRFSNSYDLERLAQTRQOIPLG 130
QY 123 IEELIOSVALRYPGSGSTRQAARSLIILOMISEAARFNPFMRARQYINSGESFLPDMY 182
Db 131 LQALTHAISFLQ-SGTDQGEIARLVIITQMASEAARFYISNRGVSIKRTGAFQPDPA 189
QY 183 MLELETSWGQOSTQVOQSTDGVE--NNPFRLGISTGNEFTLSNVADVIASLAIMLFCRDR 241
Db 190 MLELENNMNDNLSCGVQOSYODTFEPNAVTLRSVNNQPVIVDSLTHQSVAVIALMLFVC-NP 248
QY 242 PSSSDVRYMPLVIRPVLENSGAVDDVCTASPTVRI 278
Db 249 PNAQOS---PLLRISIVEES---KICSSRYEPTVRI 278
```

RESULT 6

```
C39761
abrin (clone 7.2) precursor - Indian licorice (fragment)
N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
C:Species: Abrus precatorius (Indian licorice)
C:Date: 21-Feb-1992 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
C:Accession: C39761; S14471
```

Query Match	32.1%;	Score 458;	DB 2;	Length 251;
Best Local Similarity	45.6%;	Pred. No. 8.6e-34;		
Matches 108;	Conservative 34;	Mismatches 91;	Indels 4;	Gaps 4

RESULT 7

Query Match	30.08;	Score 427;	DB 2;	Length 570;
Best Local Similarity	37.68;	Pred. No. 1.7e-30;		

3

RESULT 8

RLCSD

A;Accession: A24614
A;Molecule type: mRNA
A;Residues: 12-75, 'D', 77-550, 'R', 552-576 <LAM>

R.Yoshitake, S.: Funatsu, G.; Funatsu, M.
Agric. Biol. Chem. 42, 1267-1274, 1978
A>Title: Isolation and sequences of peptic peptides, and the complete sequence of ribonuclease A
A:Reference number: A03372
A:Accession: A03372
A:Molecule type: protein
A:Residues: 36-97, 'Q', 99-109, 'S', 111-269, 'D', 272-283, 'V', 285-288, 290-302 <XOS>
A>Note: this paper cites the others in the series providing experimental details for
R.Rarki, T.; Funatsu, G.
FEBS Lett. 191, 121-124, 1985
A>Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptophan
A:Reference number: A24010
A:Accession: A24010
A:Molecule type: protein
A:Residues: 315-383, 'PS', 386-576 <ARA>
R.Funatsu, G.; Kimura, M.; Funatsu, M.
Agric. Biol. Chem. 43, 2221-2224, 1979
A>Title: Primary structure of Ala chain of ricin D.
A:Reference number: A03374
A:Accession: A03374
A:Molecule type: protein
A:Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 527, 'E', 529-564, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F' <UN>
A>Note: this paper, one of a series, summarizes the experimental details for the detection of ricin in plant material.
R:Ready, M.P.; Kim, Y.; Robertus, J.D.

A, Reference number: A6/092; FDB: LMAN

activation activities. It is of interest

F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match 22.9%; Score 327; DB 2; Length 247;

Best Local Similarity 37.2%; Pred. No. 5.8e-22;

Matches 89; Conservative 51; Mismatches 85; Indels 14; Gaps 9;

```

OY 4 TGDEYFRITLLRDYVSSGSPSENEIPLLRQSTIPVSDAQRFEVLVELTNGGDSITRAIDY 63
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10 TSSSYGVFISNLRKALPYERKLYDIPLLR-STLPGS--GRYALHILTNYADETISVAIDV 66
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 64 TNLVYVAVQAGDQSYFLDA-PDGAERHLFTGTRS-SLPFGSYTDLERYAGH-RDQIP 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 TNYVYVAGYRAGDTSYFENEASATEAKYVEKDAKRYTLFPGNYERLQIAAGKIRENIP 126
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 121 LGIEELIQSVSALRYPGGSTRQAARSLIILQIMISEAARFNPIFWRAQYIINGESFLPD 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 127 LGLPALDSAITLFFYNNNSAASA-LMWLIQSTSEAAARYKFTEOQIGKRVD--KTFLPS 182
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 181 MYMLELETSMGQOSTQVQ--OSTDGVENNPFRLGISTGNFVTLNSVRD--VIASLAIML 235
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 183 LATISLENSWSALSQIQIASTNNGQFETPVVLIINAQNRVTITNVDAGVVTSNIALLL 241
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: March 18, 2003, 08:27:32
 Job time : 25.1332 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 18, 2003, 08:01:23 ; Search time 12.6181 Seconds
(without alignments)
917.088 Million cell updates/sec

Title: US-09-627-165E-16

Perfect score: 1425
Sequence: 1 HQTGDEXFRTILLRDYVS.....NSGAVDVTCFASPTVRIIV 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSPROT_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1119	78.5	254	1 MIA_VISAL	P81446 viscum albu
2	514	36.1	528	1 ABRA_ABRPR	P11140 abrus prec
3	481	33.8	562	1 ABRA_ABRPR	P28590 abrus prec
4	471.5	33.1	527	1 ABRB_ABRPR	P06077 abrus prec
5	425.5	29.9	576	1 RIC1_RICCO	P02879 ricinus com
6	423	29.7	564	1 AGGL_RICCO	P06750 ricinus com
7	393	27.6	563	1 NIGB_SAMNT	P35183 sambucus n
8	354	24.8	282	1 RIP2_BRIDI	P38184 bryonia dio
9	340.5	23.9	286	1 RIP2_MOMBA	P29339 momordica b
10	336	23.6	289	1 RIP1_TRIKI	P09989 trichosan
11	331.5	23.3	289	1 RIP1_TRIKI	P24478 trichosan
12	312.5	21.9	294	1 RIP1_TRIAN	P56626 trichosan
13	311	21.8	316	1 RIPG_GELMU	P33186 gelonium mu
14	308.5	21.6	286	1 RIP1_GUCFI	O96fx4 cucumis fig
15	297.5	20.9	290	1 RIP1_BRIDI	P33185 bryonia dio
16	296.5	20.8	286	1 RIP1_MOMCH	P16094 momordica c
17	286.5	20.1	277	1 RIPA_LURCY	P22851 luffa cylin
18	280	19.6	250	1 RIPA_LURCY	P22851 luffa cylin
19	249	17.5	278	1 RIPP_MIRJA	P21326 mirabilis j
20	203	14.2	313	1 RIPP_MIRJA	P10297 phyto
21	202.5	14.2	261	1 RIPP_MIRJA	P10297 phyto
22	196	13.8	299	1 RIPP_MIRJA	P23339 phyto
23	192	13.5	299	1 RIPP_MIRJA	P03464 phyto
24	188	13.2	253	1 RIPP_MIRJA	P20656 saponaria o
25	186	13.1	253	1 RIPP_MIRJA	O41391 saponaria o
26	182	12.8	292	1 RIPP_MIRJA	P27559 saponaria o
27	157.5	11.1	310	1 RIPP_MIRJA	O40772 phyto
28	148.5	10.4	293	1 RIPP_MIRJA	P24476 dianthus ca
29	143.5	10.1	296	1 RIPP_MIRJA	P27560 saponaria o
30	128	9.0	319	1 RIPP_MIRJA	P09388 bacterioph
31	120.5	8.5	315	1 RIPP_MIRJA	P08026 bacterioph
32	120.5	8.5	315	1 RIPP_MIRJA	P10149 bacterioph
33	120	8.4	280	1 RIPP_MIRJA	P22244 hordeum vul

34	113	7.9	280	1	RIP2_HORVU	P04399 hordeum vul
35	103.5	7.3	560	1	J160_HORVU	O00531 hordeum vul
36	92.5	6.5	468	1	PPAD_YEAST	P52290 saccharomyc
37	92	6.5	574	1	YYP3_CABEL	P52716 caenorhabdl
38	90.5	6.4	400	1	B3AR_MOUSE	P25962 mus musculu
39	89.5	6.3	351	1	B3AR_CAVPO	O60483 cavla porce
40	88.5	6.2	971	1	SECS_YEAST	P89102 saccharomyc
41	88	6.2	467	1	B1AR_SHEEP	O28927 ovis aries
42	87.5	6.1	400	1	B3AR_RAT	P26256 rattus norv
43	86	6.0	600	1	MGCD_SALTY	P33902 salmone
44	85.5	6.0	405	1	B3AR_CANFA	O02662 canis fami
45	85	6.0	454	1	DAT_HABIN	P44951 h diaminobu

ALIGNMENTS

RESULT 1	MLA_VISAL	STANDARD:	PRT:	254 AA.
AC	P81446,			
DT	15-DEC-1998 (rel. 37, Created)			
DT	15-DEC-1998 (rel. 37, Last sequence update)			
DT	15-JUN-2002 (rel. 41, Last annotation update)			
DE	Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (rRNA N-glycosidase) (EC 3.2.2.22).			
OS	Viscum album (European mistletoe).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Santalales; Viscaceae; Viscum.			
OX	NCBI_TaxID=3972;			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN-Subsp. album:			
RX	MEDLINE-97134581; PubMed-6960141;			
RA	Huguet Soler M., Stoeva S., Schwaborn C., Wilhelm S., Stiefel T.,			
RA	Voelter W.;			
RT	"Complete amino acid sequence of the A chain of mistletoe lectin I.";			
RL	FEBS Lett. 399:153-157(1996).			
CC	-1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY INTO THE CELL OF THE A CHAIN. B CHAINS ARE ALSO RESPONSIBLE FOR CELL AGGLOUTINATION (LECTIN ACTIVITY).			
CC	-1- CATALYTIC ACTIVITY: Endonuclease of the N-glycosidic bond at one specific adenosine on the 28S rRNA.			
CC	-1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.			
CC	-1- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA AND NON-GLYCOSYLATED FORM MLA'.			
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.			
DR	HSP: P1140; IABR.			
DR	Interpro: IPR001574; RIP.			
DR	Pfam: PF00161; RIP.			
DR	PRINTS: PR00396; SHIGARICIN.			
DR	PROSITE, PS00275: SHIGARICIN; FALSE_NEG.			
KW	Plant defense; Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.			
FT	ACT_SITE 165 .. 165			
FT	CARBOHYD 112 .. 112			
FT	VARIANT 15 .. 15			
FT	VARIANT 112 .. 112			
FT	VARIANT 66 .. 66			
FT	VARIANT 112 .. 112			
FT	VARIANT 116 .. 116			
FT	VARIANT 133 .. 134			
FT	VARIANT 140 .. 140			
FT	VARIANT 144 .. 144			
FT	VARIANT 151 .. 151			
FT	VARIANT 179 .. 179			
FT	VARIANT 184 .. 184			
FT	VARIANT 190 .. 190			
FT	VARIANT 218 .. 218			
FT	VARIANT 223 .. 224			
FT	VARIANT			

FT VARIANT 231 231 T -> S (IN M.L.A.).
 FT VARIANT 235 235 D -> S (IN M.L.A.).
 SQ SEQUENCE 254 AA; 28478 MW; 53BAF98D3D0FF667 CRC64;
 Query Match 78.5%; Score 1119; DB 1; Length 254;
 Best Local Similarity 89.0%; Pred. No. 3.4e-93;
 Matches 218; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 HQTGDEFFRFTLLRDYVSSGSGSNELPLRSTIVSDAQRFLVETLNGGDSITAA 60
 D 10 HQTGEFFRFTLLRDYVSSGSGSNELPLRSTIVSDAQRFLVETLNGGDSITAA 69
 QY 61 IDVTNLVVAOAGDQSYFLRDADPGARHLFTGTSSSLPFGSGYDLEERYGHDQIP 120
 D 70 IDVTNLVVAOAGDQSYFLRDADPGARHLFTGTSSSLPFGSGYDLEERYGHDQIP 129
 QY 121 LGTEELIOSVALRPGSGTRQARSLLILQIMISEARFNPFWARQYINGSEFLPD 180
 D 130 LGIDOLIOSTALRPGSGTRQARSLLILQIMISEARFNPILMRQYINGSEFLPD 189
 QY 181 MYMLEETSMGQSTQVQSTQGVNPFRLGISTGNFVTLNVRDYATSLAIFVCRD 240
 D 190 MYMLEETSMGQSTQVQSTQGVNPFRLAIPGNEFVTLNVRDYATSLAIFVCGE 249
 QY 241 RPSSS 245
 D 250 RPSSS 254

RESULT 2
 ABRA_ABRP STANDARD; PRT; 528 AA.
 ID ABRA_ABRP
 AC P11140; F28589; (Rel. 11, Created)
 DT 01-JUN-1989 (Rel. 29, Last annotation update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Abria-a precursor [Contains: Abria-a A chain (rRNA N-glycosidase)
 (EC 3.2.2.22); Abria-a B chain].
 OS Abria-a precursor (Indian lilycote) (Crab's eye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 NC NCBITaxID=3816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93132798; PubMed=8421313;
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
 RT "Primary structure of three distinct isoabrians determined by cDNA
 sequencing. Conservation and significance.";
 RL J. Mol. Biol. 229:263-267(1993).
 RN [2]
 RP SEQUENCE OF 1-251.
 RC Tissue=Seed.
 RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
 RT "The complete amino acid sequence of the A-chain of abria-a, a toxic
 protein from the seeds of Abrus precatorius.";
 RL Agric. Biol. Chem. 52:1095-1097(1988).
 RN [3]
 RP SEQUENCE OF 1-251 FROM N.A.
 RC Tissue=Leaf.
 RX MEDLINE=91201329; PubMed=2016300;
 RA Evensen G., Mathiesen A., Sundan A.;
 RT "Direct molecular cloning and expression of two distinct abria
 A-chains.";
 RL J. Biol. Chem. 266:6848-6852(1991).
 RN [4]
 RP SEQUENCE OF 262-528.
 RX MEDLINE=92371656; PubMed=1505674;
 RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
 RT "The complete primary structure of abria-a B chain.";
 RL FEBS Lett. 309:115-118(1992).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).

RX MEDLINE=95333188; PubMed=7608980;
 RA Tahirou T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
 RT "Crystal structure of abria-a at 2.14 A.";
 RL J. Mol. Biol. 250:354-367(1995).
 CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.
 CC -1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
 FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
 PRECEDES ENDOCYTOSIS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
 INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M98344; AAA32624.1; ALF_INIT.
 CC EMBL: X54872; -; NOT_ANNOTATED_CDS.
 CC PIR: S32429; T2LSA.
 CC PIR: S24133; S24133.
 CC PDB: 1ABR; 07-FEB-95.
 CC InterPro: IPR001574; RIP.
 CC InterPro: IPR000772; Ricin_B_lectin.
 CC Pfam: PF00652; Ricin_B_lectin; 6.
 CC Pfam: PF00652; SHIGARGICIN.
 CC PRINTS: PR00396; SHIGARGICIN.
 CC SMART: SM00458; RICIN; 2.
 CC PROSITE: PS00275; SHIGA_RICIN; 1.
 CC PROSITE: PS00275; SHIGA_RICIN; 1.
 CC KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KW Glycoprotein; Lactin; 3D-structure.
 KW CHAIN 1 251
 FT PEPTIDE 252
 FT CHAIN 262 528
 FT DOMAIN 273 400
 FT DOMAIN 403 527
 FT REPEAT 283 325
 FT REPEAT 326 366
 FT REPEAT 369 401
 FT REPEAT 414 449
 FT REPEAT 453 492
 FT REPEAT 495 528
 FT REPEAT 164 164
 FT ACCT_SITE 164 269
 FT DISULFID 247 269
 FT DISULFID 286 305
 FT DISULFID 329 346
 FT DISULFID 417 430
 FT DISULFID 456 473
 FT MOD_RES 1 1
 FT CARBOHYD 361 361
 FT CARBOHYD 401 401
 FT CONFLICT 202 202
 FT CONFLICT 298 298
 FT CONFLICT 427 427
 FT CONFLICT 467 467
 FT CONFLICT 483 483
 SQ SEQUENCE 528 AA; 59243 MW; A1F7B6CD5B9A827 CRC64;
 Query Match 36.1%; Score 514; DB 1; Length 528;
 Best Local Similarity 42.5%; Pred. No. 1.8e-38;
 Matches 122; Conservative 45; Mismatches 88; Indels 32; Gaps 8;

RX MEDLINE-93132798; PubMed-8421313;
 RA Hung C.-H., Lee M.-C., Lee J.-C., Lin J.-Y.;
 RT "Primary structure of three distinct isoforms determined by cDNA
 RT sequencing. Conservation and significance.";
 RL J. Mol. Biol. 229:263-267(1993).
 RN [2]
 RP SEQUENCE OF 260-527.
 RC TISSUE-Seed.
 RA MEDLINE-93169023; PubMed-7763422;
 RA Kimura M., Sumizawa T., Funatsu G.;
 RT "The complete amino acid sequences of the B-chains of abrin-a and
 RL abrin-b, toxic proteins from the seeds of Abrus precatorius.";
 RL Biosci. Biotechnol. Biochem. 57:166-169(1993).
 CC -I- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.
 CC -I- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
 CC PRECEDES ENDOCYTOSIS.
 CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -I- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -I- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
 CC INACTIVATING PROTEIN FAMILY, TYPE 2 RIP SUBFAMILY.
 CC -I- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; M98345; AAA32625.1; -
 DR HSSP; P11140; IABR.
 DR InterPro; IPR001574; RIP.
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR Pfam; PF00161; RIP; 1.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 DR Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KW Glycoprotein; Lectin.
 KM CHAIN 1 250 ABRIN-B A CHAIN.
 FT PEPTIDE 251 260 LINKER PEPTIDE.
 FT CHAIN 261 527 ABRIN-B B CHAIN.
 FT DOMAIN 272 399 RICIN B-TYPE LECTIN 1.
 FT DOMAIN 402 526 RICIN B-TYPE LECTIN 2.
 FT REPEAT 282 324 1-ALPHA.
 FT REPEAT 325 365 1-BETA.
 FT REPEAT 368 400 1-GAMMA.
 FT REPEAT 413 448 2-ALPHA.
 FT REPEAT 452 491 2-BETA.
 FT REPEAT 494 527 2-GAMMA.
 FT ACT_SITE 163 163 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 246 268 BY SIMILARITY.
 FT DISULFID 285 304 BY SIMILARITY.
 FT DISULFID 328 345 BY SIMILARITY.
 FT DISULFID 416 429 BY SIMILARITY.
 FT DISULFID 455 472 BY SIMILARITY.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY
 FT SIMILARITY).
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 282 282 N -> D (IN REF. 2).
 FT CONFLICT 291 291 D -> N (IN REF. 2).

FT CONFLICT 350 351 AE -> PQ (IN REF. 2).
 FT CONFLICT 378 378 S -> N (IN REF. 2).
 FT CONFLICT 426 426 Y -> M (IN REF. 2).
 FT CONFLICT 428 428 X -> D (IN REF. 2).
 FT CONFLICT 431 431 N -> S (IN REF. 2).
 FT CONFLICT 484 484 R -> K (IN REF. 2).
 FT CONFLICT 491 491 N -> S (IN REF. 2).
 FT CONFLICT 493 493 H -> Y (IN REF. 2).
 FT CONFLICT 502 502 R -> Q (IN REF. 2).
 FT CONFLICT 509 509 E -> Q (IN REF. 2).
 FT CONFLICT 513 513 H -> W (IN REF. 2).
 FT CONFLICT 516 516 H -> T (IN REF. 2).
 SQ SEQUENCE 527 AA; 59114 MW; 3253AE490CE9494A CRC64;
 Query Match 33.1%; Score 471.5; DB 1; Length 527;
 Best Local Similarity 42.6%; Pred. No. 1,2e-34;
 Matches 118; Conservative 46; Mismatches 100; Indels 13; Gaps 8;
 QY 4 TGDEYFETLLADYVSSGSFSENEIPLRSTIPVSAORFVVELTNGSGSTPAIDV 63
 DB 13 TSQSYKOFITALKQRL-TGGLHGIPVLPDPT-TLQERNRIVVELSNDTSTIAGIDV 70
 QY 64 TNLVVAVQAGDSYFLRDAPGAERHLFTGTRSSLPFTGTDLERYAGH-RQQIPLG 122
 DB 71 SNAVVAVRAGNSYFLRDAPPSASRYLFTGTQOYSLRFNGSYIDLERLARQTRQIPLG 130
 QY 123 IEEILQSVSALRYPGSGSTGAARSLLIQLQMSHAFNFIPIRARQYINSGESFLPDMY 182
 DB 121 LQALRAHISFLQ-SGTDDIELARTLIVLIQMSAPARYRISYVGVGSIYRTWTAQPDA 189
 QY 183 MLEETSGWGQSGVQVQSGVFNPNPRL-GISGNGVTLSTNRYDIASIALMLFVCHDR 241
 DB 190 MISLENMMDLGSQVQSGVQDFPNNAVILRSVNQPIYVSLTHQSVANLALMLFVC-NP 248
 QY 242 PSSSDVRYPLVIRPVLENSGAVDDVYTASEPVRI 278
 DB 249 PNAQNS---PLLIIRSIKVS---KICSSRYEPTVRI 278
 RESULT 5
 RICIN_RICCO STANDARD; PRT; 576 AA.
 AC P02879; P02880;
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ricin precursor [contains: Ricin A chain (rRNA N-glycosidase)
 DE (EC 3.2.2.22); Ricin B chain].
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Malpighiales; Euphorbiaceae; Ricinus.
 NCBI_TaxID=3988;
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86067214; PubMed-2999712;
 RA Halling K.C., Halling A.C., Murray E.F., Ladin B.F., Houston L.L.,
 RA Weaver R.F.;
 RT "Genomic cloning and characterization of a ricin gene from Ricinus
 RT communis.";
 RL Nucleic Acids Res. 13:8019-8033(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92163016; PubMed-1371405;
 RA Tregear J.W., Roberts L.M.;
 RT "The ricin gene family of Ricinus communis: cloning of a functional
 RT ricin gene and three lectin pseudogenes.";
 RL Plant Mol. Biol. 18:515-525(1992).
 RN [3]
 RP SEQUENCE OF 12-576 FROM N.A.
 RX MEDLINE-85179479; PubMed-3838723;
 RA Lamb A., Roberts L.M., Lord J.M.;
 RT "Nucleotide sequence of cloned cDNA coding for preproricin.";

RL Eur. J. Biochem. 148:265-270(1985).
 RN
 RN SEQUENCE OF 36-302.
 RA Yoshitake S., Funatsu G., Funatsu M.;
 RT "Isolation and sequences of peptic peptides, and the complete
 RT sequence of the chain of ricin-D.";
 RL Agric. Biol. Chem. 42:1267-1274(1978).
 RN
 RN SEQUENCE OF 315-576.
 RA Funatsu G., Kimura M., Funatsu M.;
 RT "Primary structure of the chain of ricin D.";
 RL Agric. Biol. Chem. 43:2221-2224(1979).
 RN
 RN CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
 RA MEDLINE-90344223; PubMed-1568517;
 RA Kimura Y., Kusuhoku H., Tada M., Takagi S., Funatsu G.;
 RT "Structural analyses of sugar chains from ricin A-chain variant.";
 RL Agric. Biol. Chem. 54:157-162(1990).
 RN
 RN REVIEW.
 RA MEDLINE-21480122; PubMed-11595634;
 RA Olanes S., Kozlov J.V.;
 RT "Ricin";
 RL Toxicol. 39:1723-1728(2001).
 RN
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RA MEDLINE-87163983; PubMed-358597;
 RA Montfort W., Villalanza J.E., Monzingo A.F., Ernst S.R., Katzin B.,
 RA Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.;
 RT "The three-dimensional structure of ricin at 2.8 A.";
 RL J. Biol. Chem. 262:5398-5403(1987).
 RN
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
 RA MEDLINE-91352004; PubMed-1881881;
 RA Katzin B.J., Collins E.J., Robertus J.D.;
 RT "Structure of ricin A-chain at 2.5 A.";
 RL Proteins 10:251-259(1991).
 RN
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN.
 RA MEDLINE-91352005; PubMed-1881882;
 RA Rutenber E., Robertus J.D.;
 RT "Structure of ricin B-chain at 2.5-A resolution.";
 RL Proteins 10:260-269(1991).
 RN
 RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN.
 RA MEDLINE-95082010; PubMed-7990130;
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J.,
 RA Baupitt R.A.;
 RT "X-ray structure of recombinant ricin A-chain at 1.8-A resolution.";
 RL J. Mol. Biol. 244:410-422(1994).
 RN
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215.
 RA MEDLINE-96374222; PubMed-8780513;
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M.,
 RA Molina-Sylth M.C., Robertus J.D.;
 RT "Structure and activity of an active site substitution of ricin A
 chain.";
 RL Biochemistry 35:11098-11103(1996).
 RN
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN.
 RA MEDLINE-97240820; PubMed-9086280;
 RA Yan X., Hollis T., Sylth M., Day P., Monzingo A.F., Milne G.W.,
 RA Robertus J.D.;
 RT "Structure-based identification of a ricin inhibitor.";
 RL J. Mol. Biol. 266:1043-1049(1997).
 RN
 RN MUTAGENESIS.
 RA MEDLINE-93165632; PubMed-1287657;
 RA Kin Y., Robertus J.D.;
 RT "Analysis of several key active site residues of ricin A chain by
 RT mutagenesis and X-ray crystallography.";
 RL Protein Eng. 5:775-779(1992).
 CC
 CC -I- FUNCTION: Ricin is highly toxic to animal cells and to a less

CC extent to plant cells. The A chain is responsible for inhibiting
 CC protein synthesis through the catalytic inactivation of 60S
 CC ribosomal subunits. It acts as a glycosidase that removes a
 CC specific adenine residue from an exposed loop of 28S ribosomal
 CC RNA. As this loop is involved in the binding of elongation
 CC factors, the modified ribosomes are unable to support protein
 CC synthesis. The A chain can inactivate a few thousand ribosomes
 CC per minute, thus inactivating them faster than the cell can make
 CC new ones. A single A-chain molecule can therefore kill an animal
 CC cell. The B chain binds to cell receptors and facilitates the
 CC entry into the cell of the A chain; B chains are also responsible
 CC for cell agglutination (lectin activity). It binds to beta-D-
 CC galactopyranoside moieties.
 CC
 CC -I- CATALYTIC ACTIVITY: Endoglycolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC
 CC -I- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC
 CC -I- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC
 CC -I- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE
 CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271
 CC
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC
 CC -I- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS
 CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3).
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-slb.ch/announce/>
 CC or send an email to license@slb-slb.ch).
 CC -----
 CC
 CC EMBL: X03179; CAA26339.1; -;
 CC EMBL: X52908; CAA37095.1; -;
 CC EMBL: X02388; CAA26230.1; -;
 CC EMBL: A12892; CAA01058.1; -;
 CC PIR: A24041; RLCSD.
 CC PDB: 2AA1; 3I-JAN-94.
 CC PDB: 1A9G; 3I-JAN-94.
 CC PDB: 1FMP; 3I-OCT-93.
 CC PDB: 1IFS; 14-JAN-98.
 CC PDB: 1IFS; 14-JAN-98.
 CC PDB: 1IFT; 14-JAN-98.
 CC PDB: 1IFU; 14-JAN-98.
 CC PDB: 1RTC; 3I-OCT-93.
 CC PDB: 1OBS; 16-JUN-97.
 CC PDB: 1OBT; 16-JUN-97.
 CC PDB: 1BR5; 02-SEP-98.
 CC PDB: 1BR6; 02-SEP-98.
 CC
 CC GLOSUIDB: P02879; -;
 CC InterPro: IPR001574; RIP.
 CC InterPro: IPR00772; Ricin_B_lectin.
 CC Pfam: PF00161; RIP. 1.
 CC Pfam: PF00652; Ricin_B_lectin; 6.
 CC PRINTS: PR00396; SHIGARICIN.
 CC SMART: SM00438; RICIN; 2.
 CC PROSITE: PS00231; RICIN_B_LECTIN; 2.
 CC PROSITE: PS00275; SHIGA_RICIN; 1.
 CC Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 CC Glycoprotein; Lectin; Signal; 3D-structure.
 CC
 CC SIGNAL 1 35
 CC CHAIN 36 302 RICIN A CHAIN.
 CC PEPTIDE 303 314 LINKER PEPTIDE.
 CC CHAIN 315 576 RICIN B CHAIN.
 CC DOMAIN 321 448 RICIN B-TYPE LECTIN 1.
 CC DOMAIN 451 575 RICIN B-TYPE LECTIN 2.
 CC REPEAT 331 373 1-ALPHA.
 CC REPEAT 374 414 1-BETA.
 CC REPEAT 417 449 1-GAMMA.
 CC REPEAT 462 497 2-ALPHA.
 CC REPEAT 501 540 2-BETA.
 CC REPEAT 543 570 2-GAMMA.

```

FT ACT SITE 212 212
FT DISULFID 294 318 INTERCHAIN.
FT DISULFID 334 353
FT DISULFID 377 394
FT DISULFID 465 478
FT DISULFID 504 521
FT CARBOHYD 45 45
FT CARBOHYD 271 271
FT CARBOHYD 271 271
FT CARBOHYD 409 409
FT CARBOHYD 449 449
FT CONFLICT 76 76
FT CONFLICT 551 551
FT STRAND 43 47
FT TURN 49 50
FT TURN 53 67
FT HELIX 73 74
FT STRAND 75 76
FT TURN 77 79
FT STRAND 88 90

Query Match 29.9%; Score 425.5; DB 1: Length 576;
Best Local Similarity 40.6%; Pred. No. 1,7e-30;
Matches 117; Conservative 45; Mismatches 103; Indels 23; Gaps 11;

OY 4 TGDEYFRFTLLRDYVSSGS-FSNEIPIL-RQSTIPVSDAORFVLELTVNOGDSITTAI 61
DB 52 TVQSYTNPIRNVGRGLTGVADRHPIVLPNRYGLPIN-QRFLIVELSHNAELSTITAL 109
OY 62 DVTNLYVAYAGDQGYFLDAPDAE-----RHLEFGT-RSSLPTGSDYDLDEYGH 115
DB 110 DVTNLYVAYAGDQGYFLDAPDAE-----RHLEFGT-RSSLPTGSDYDLDEYGH 167
OY 116 -RDQIPGIEELIQSVSLRY---PGGSTRQAQNSLILTIQMISEARENPJFWARROYI 171
DB 168 LRENIENLGNGLFEAISALYYSTGTQDLPILARSFIICQIMISEARFQYIEGEMRI 227
OY 172 NSGESFLPDYMLELSTMGQOSTOVQSTGVENPFRLGISTGNFVTLNRYVIA 231
DB 228 RYNRSAPDSVITLSEMSGRSLTAIOESNOCASFASPIQORRNGSFYDVSTILPII 287
OY 232 AIMLFVCRDRSSSDYRWMPVIRPYLENSGAVDVCYTASEPVRIV 279
DB 268 ALMYRCAPRPSSO-----PSLIRPVYPMFN--DV-CMDEPIVIRIV 328

RESULT 6
AGGL_RICCO STANDARD; PRT; 564 AA.
ID AGGL_RICCO
AC P06750;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Agglutinin precursor (PCA) [Contains: Agglutinin A chain (rRNA N-
glycosidase) (EC 3.2.2.22); Agglutinin B chain].
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids I; Malpighiales; Euphorbiaceae; Ricinus.
OK NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86059449; PubMed=2999130;
RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;
RT "The primary sequence of Ricinus communis agglutinin. Comparison with
RT ricin."
RL J. Biol. Chem. 260:15682-15686(1985).
RN [2]
RP SEQUENCE OF 303-564.
RC TISUP-Seed;
RA Ataki T., Yoshioka Y., Funatsu G.;
```

```

RT "The complete amino acid sequence of the B-chain of the Ricinus
RT communis agglutinin isolated from large-grain castor bean seeds.";
RT Biochim. Biophys. Acta 872:277-285(1986).
RN [3]
RP SEQUENCE OF 303-337.
RX MEDLINE=80178723; PubMed=6768555;
RA Lin T.T.-S., Li S.S.-L.;
RT "Purification and physicochemical properties of ricin and
RT agglutinins from Ricinus communis."
RL Eur. J. Biochem. 105:453-459(1980).
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -I- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-
CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
CC -I- SIMILARITY: CONTAINS 2 RICIN B-TYPE LECTIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12089; AAA33869.1; -
DR EMBL; S40368; AAB22584.1; -
DR PIR; A24261; RLC5AG.
DR HSSP; P02879; 1BR6.
DR GlycosylatedB; P06750; -.
DR Interpro: IPR001574; RIP.
DR Interpro: IPR000772; Ricin_B_lectin.
DR Pfam; PR00161; RIP; 1.
DR Pfam; PR00652; Ricin_B_lectin; 6.
DR SMART; SM00458; RICIN; 2.
DR SMART; PR00396; SHIGARICIN.
DR PROSITE; PS50231; RICIN_B_LLECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Hydrolyase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Signal.
FT SIGNAL 1 24
FT CHAIN 25 290 POTENTIAL.
FT PROPEP 291 302 AGGLUTININ A CHAIN.
FT CHAIN 303 564 LINKER PEPTIDE.
FT DOMAIN 309 436 AGGLUTININ B CHAIN.
FT DOMAIN 439 563 RICIN B-TYPE LECTIN 1.
FT REPEAT 319 361 RICIN B-TYPE LECTIN 2.
FT REPEAT 362 402 1-ALPHA.
FT REPEAT 403 427 1-BETA.
FT REPEAT 428 457 1-GAMMA.
FT REPEAT 458 485 2-ALPHA.
FT REPEAT 489 528 2-BETA.
FT REPEAT 531 558 2-GAMMA.
FT ACT_SITE 200 200 BY SIMILARITY.
FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).
FT DISULFID 322 341 BY SIMILARITY.
FT DISULFID 365 382 BY SIMILARITY.
FT DISULFID 453 466 BY SIMILARITY.
FT DISULFID 492 509 BY SIMILARITY.
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 331 331 F -> T (IN REF. 2).
FT CONFLICT 362 362 N -> D (IN REF. 2).
FT CONFLICT 374 374 R -> G (IN REF. 2).
FT CONFLICT 404 404 F -> T (IN REF. 2).
FT CONFLICT 552 552 F -> V (IN REF. 2).
SQ SEQUENCE 564 AA; 62851 MW; DA55F2AV2F60959 CRC64;
```

```

Query Match 29.7%; Score 423; DB 1: Length 564;
Best Local Similarity 39.7%; Pred. No. 2.8e-30;
Matches 114; Conservative 45; Mismatches 106; Indels 22; Gaps 10;
```

```
OY 4 TGDEYFRFTLLRDYVSSGS-FSNEIPIL-RQSTIPVSDAORFVLELTVNOGDSITTAI 61
```


DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (BD2).

OS Bryonia dioica (red bryony).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eucosids I; Cucurbitales; Cucurbitaceae; Bryonia.

NCBI_TaxID=3652;

OX NCBI_TaxID=3652;

RN SEQUENCE FROM N.A.

RA Siegal C.B., Gawlak S.L., Marguardt H.;

RT "Bryodin 2, a ribosome-inactivating protein isolated from the plant Bryonia dioica."

RL Patent number US5597569, 28-JAN-1997.

RP [2]

RP SEQUENCE OF 22-42.

RC TISSUE-Root;

RC MEDLINE=95151812; PubMed=7849072;

RX Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B., Marguardt H.;

RA "Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunoconjugates."

RT Bioconj. Chem. 5:423-429(1994).

RL -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.

CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

CC TYPE 1 RIP SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: I34238; NOT_ANNOTATED_CDS.

DR HSSP: P09989; IMRJ.

DR InterPro: IPR001574; RIP.

DR Pfam: PF00161; RIP; 1.

DR PRINTS: PR00396; SHIGARICIN.

DR PROSITE: PS00275; SHIGA_RICIN; 1.

DR Plant defense: Protein synthesis inhibitor; Hydrolase; Toxin; Signal.

KW Multigene family; Glycoprotein; Signal.

FT SIGNAL 1 21

FT CHAIN 22 282

FT ACT_SITE 183 183

FT CARBOHD 25 25

FT SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;

SO Query Match 24.8%; Score 354; DB 1; Length 282;

Best Local Similarity 39.3%; Pred. No. 1.7e-24;

Matches 95; Conservative 37; Mismatches 92; Indels 18; Gaps 9;

QY 4 TGDYFRFTLLRDVYSSGSFNSN---EIPLRQSTIPVSDAQRVAVLVELTNOGGDSTIAAD 62

DB 32 TGAFTYFTLRNRTKLTGTPRVVDIPLVLRNAAAGLA---RFOVLVLTNNNGSVTVAD 88

QY 63 VTMVYVAYOAGDQSYFLADPDGAEHLFTGTRSSLPFTGSDTDEYAGH--RDQIP 120

DB 89 VVAVVYVAYRAGTAFELADASTEANNVLFAGINHVRLPYGNGDGLETAAGRISRENIE 148

QY 121 LGIEELIOSVSL--RYPGSTRQAQRSLIILQMISEAFRNPIFWRARQYINGSESL 178

DB 149 LGSEISSAIGNMFRRNPSTV---PRATIVITQVSEARPKYIEQVSE--NVGTRFK 203

QY 179 PDWYMLELETSMGQOSTOV--QSTDGVFNPNFRIGT--SIGNFVTLTNSVDYIASLAIM 234

DB 204 PDPAFLSLQANMSLSKQIPLAQNQSGKFRNPVDLTKPTGERFOVTVNDSVYKGNIKLL 262

QY 235 LF 236

DB 263 LY 264

1:

RESULT 9

RIP2_MOMBA STANDARD; PRT; 286 AA.

ID RIP2_MOMBA

AC P29339;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ribosome-inactivating protein momordin II precursor (rRNA N-glycosidase) (EC 3.2.2.22).

DE Momordica balsamina (Bitter melon) (Balsam pear).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eucosids I; Cucurbitales; Cucurbitaceae; Momordica.

NCBI_TaxID=3672;

OX NCBI_TaxID=3672;

RN SEQUENCE FROM N.A.

RP TISSUE-Seed;

RC MEDLINE=93027170; PubMed=1408771;

RX Orliga M., Better M.;

RA "Momordin II, a ribosome inactivating protein from Momordica balsamina, is homologous to other plant proteins."

RT Nucleic Acids Res. 20:4662-4662(1992).

RL -1- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.

CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.

CC TYPE 1 RIP SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: Z12175; CAA78166.1; -

DR PIR: S25560; S25560.

DR HSSP: P09989; IMRJ.

DR InterPro: IPR001574; RIP.

DR Pfam: PF00161; RIP; 1.

DR PRINTS: PR00396; SHIGARICIN.

DR PROSITE: PS00275; SHIGA_RICIN; 1.

DR Plant defense: Protein synthesis inhibitor; Hydrolase; Toxin; Signal.

KW Multigene family; Glycoprotein; Signal.

FT SIGNAL 1 23

FT CHAIN 24 286

FT ACT_SITE 181 181

FT SEQUENCE 286 AA; 32031 MW; 3B89FF1AE6825986 CRC64;

SO Query Match 23.9%; Score 340.5; DB 1; Length 286;

Best Local Similarity 35.2%; Pred. No. 2.8e-23;

Matches 89; Conservative 53; Mismatches 92; Indels 19; Gaps 9;

QY 4 TGDYFRFTLLRDVYSSGSFNSN---EIPLRQSTIPVSDAQRVAVLVELTNOGGDSTIAAD 60

DB 33 TAKYTFEI---EDFRATLPFSHKYIDLPPLAST---ISDRRLILDLSVAETISVA 86

QY 61 IDVTNLYVYVAYOAGDQSYFLADPDGAEHLFTGTRSSLPFTGSDTDEYAGH--RDQ 118

DB 87 IDVTNLYVYVAYRAGTAFELADASTEANNVLFAGINHVRLPYGNGDGLETAAGRISRENIE 145

QY 119 IPIGIEELIOSVSL--RYPGSTRQAQRSLIILQMISEAFRNPIFWRARQYINGSESL 178

DB 146 IDLGLPALSSAITLTFYVNAQSAASA--LIVLQTTAEARKYIERHVAVYV--ATNFK 201

QY 179 PDWYMLELETSMGQOSTOV--QOSTDGVFNPNFRIGT--SIGNFVTLTNSVDYIASLAIM 236

DB 202 PMLATISLENQMSALSQIPLAQNQSGKFRNPVDLTKPTGERFOVTVNDSVYKGNIKLL 261

QY 237 VCDRPSDDVRY 249
DB 262 L-NSRASTADENF 273

RESULT 10

ID RIPS_TRIKI

AC P09989; STANDARD; PRT; 289 AA.

DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
OS Ribosome-inactivating protein alpha-trichosanthin precursor
OC (rRNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS)
OC Trichosanthin Kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OC NCBI_Taxid=3677;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz;
RX MEDLINE=91153657; PubMed=1999291;
RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;
RT "Cloning of trichosanthin cDNA and its expression in Escherichia
RT coli.";
RL Gene 97:267-272(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz; TISSUE=leaf;
RX MEDLINE=90256790; PubMed=2341400;
RA Chow T., Feldman R.A., Lovett M., Platak M.;
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a
RT type I ribosome-inactivating protein.";
RL J. Biol. Chem. 265:8670-8674(1990).
RN [3]
RP SEQUENCE OF 24-270.
RC STRAIN=Maximowicz; TISSUE=tuberos root;
RX MEDLINE=90256789; PubMed=2341399;
RA Collins E.J., Robertus J.D., Lopest M., Stone K.L., Williams K.R.,
RT Wu P., Hwang K., Platak M.;
RT "Primary amino acid sequence of alpha-trichosanthin and molecular
RT models for abrin A-chain and alpha-trichosanthin.";
RL J. Biol. Chem. 265:8665-8669(1990).
RN [4]
RP SEQUENCE OF 24-270.
RC TISSUE=tuberos root;
RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,
RT Tian G.Y., Ni C.Z.;
RT "Scientific evaluation of Tian Hua Fen (THEF): history, chemistry and
RT application.";
RL Pure Appl. Chem. 58:789-798(1986).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
RX MEDLINE=94344957; PubMed=8066085;
RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
RT "Structure of trichosanthin at 1.88-A resolution.";
RL Proteins 19:4-13(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX MEDLINE=95344383; PubMed=7619070;
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RT dephasing mechanism of two ribosome-inactivating proteins.";
RL Biochem. J. 309:285-298(1995).
RN [7]
RP FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS
CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT IS
CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: M34858; AAA34207.1;
DR EMBL: J05434; AAA34206.1;
DR PIR: J70003; RUTZ.
DR PIR: J70566; J70566.
DR PIR: A36273; A36273.
DR PIR: A36274; A36274.
DR PDB: 1MRJ; 07-FEB-95.
DR PDB: 1MRK; 07-FEB-95.
DR PDB: 1TCS; 10-JUL-95.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; 1.
KW Plant defense; antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 270
FT PROPEP 271 289
FT ACT_SITE 183 183
FT CONFLICT 57 60
FT CONFLICT 82 84
FT CONFLICT 87 87
FT CONFLICT 92 92
FT CONFLICT 143 144
FT CONFLICT 196 196
FT CONFLICT 215 216
FT CONFLICT 231 231
FT CONFLICT 234 234
FT CONFLICT 246 246
FT CONFLICT 247 247
SQ
Query Match 289 AA; 31676 MW; SCE09BB630575BB9 CRC64;
Best Local Similarity 33.9%; Pred. NO. 7.2e-23;
Matches 93; Conservative 56; Mismatches 95; Indels 30; Gaps 9;

4 TGDYFRFTLLRDYVSSGFSNEIPLRSTIPSDAORFVLELNOGDSITAIIV 63
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
33 TSSYGVFTISNRKALPNERKLYDIPLLR--QKRALHLTNVADETISVAIDV 89
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
64 TNLVYAYQAGDSYFLDA-PDGERHLFTGTRR-SLEPTGSYDLERYAGH-RDQIP 120
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
90 TNYIMNGYRAGDTSYFENASATEAKYVRKATYLTYSNGYERLDTAGKIRENIP 149
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
121 LGIEELIOSVSLARKYPGSTRQAARSLIILOMISEARFNPFWBARQYINSGESFLPD 180
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
150 LGLPALDSAITTLTFYNNANSASA--LMLVLIQISEAKRKFLTEQIGKRV--TFLPS 205
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
181 MYLELETSMGQOSNOVO--QSTDGVPNNPFRGISTGNFVTLNRYDVIAIAIMLFC 238
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
206 LAILLEMSWALSKEQOIASTINGOPESPVLINONORVITTVNDAGVTSNIALLLN 265
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
239 RDRPSSSVRYKPLVIRPVLENGAVDVTCTAS 272
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
266 RNNMAAND-----DDVPMQOS 281
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
RIBOSOME-INACTIVATING PROTEIN ALPHA-
TRICHOSANTHIN.
MISSING IN MATURE PROTEIN.
BY SIMILARITY.
IPL -> LPL (IN REF. 4).
MISSING (IN REF. 4).
I -> L (IN REF. 4).
V -> VDAGLPNAVL (IN REF. 4).
KI -> GL (IN REF. 4).
K -> S (IN REF. 4).
Q -> LML (IN REF. 4).
S -> T (IN REF. 4).
S -> T (IN REF. 2).
MISSING (IN REF. 4).
T -> M (IN REF. 2).
247

RESULT 11
ID RIPS_TRIKI
AC P24478; STANDARD; PRT; 289 AA.
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Ribosome-inactivating protein karasurin precursor (rRNA
N-glycosidase) (EC 3.2.2.22).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
NCBI_TaxID=3677;
RN NCBI_TaxID=3677;
RP SEQUENCE FROM N.A.
RC TISSUE=Root tuber;
RX MEDLINE=97356562; PubMed=9212998;
RA Mizukami H., Iida K., Kondo T., Ogihara Y.;
RT "Cloning and bacterial expression of a gene encoding ribosome-
inactivating proteins, karasurin-A and karasurin-C, from Trichosanthes
kirilowii var. japonica.";
RT Kirilowii var. japonica.";
RL Biol. Pharm. Bull. 20:711-713(1997).
RN [2].
RP SEQUENCE OF 24-270.
RX MEDLINE=92005921; PubMed=1914000;
RA Toyokawa S., Takeda T., Kato Y., Wakabayashi K., Ogihara Y.;
RT "The complete amino acid sequence of an abortifacient protein,
karasurin.";
RT Chem. Bull. 39:1244-1249(1991).
RL -1- FUNCTION: ABORTION-INDUCING PROTEIN. IT INACTIVATES EUKARYOTIC
60S RIBOSOMAL SUBUNITS.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: AB000666; BAA21786.1; -.
DR PIR: J00393; J00393.
DR HSSP: P09989; 1MRJ.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN.1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 270 KARASURIN-C.
FT PROPEP 24 270 KARASURIN-A.
FT ACT_SITE 183 289 REMOVED IN MATURE FORM.
FT ACT_SITE 183 183 BY SIMILARITY
SQ SEQUENCE 289 AA; 31704 MW; 883D3E3242887B26 CRC64;

Query Match 23.3%; Score 331.5; DB 1; Length 289;
Best Local Similarity 34.7%; Pred. No. 1.8e-22;
Matches 93; Conservative 51; Mismatches 93; Indels 31; Gaps 9;

OY 239 RDRPSSDVRYMPLVIRPLENSGAVD 266
Db 266 RN-----NMAAIDD 274

RESULT 12
ID RIP1_TRIAN STANDARD; PRT; 294 AA.
AC P56626; Q9ZQY7.
DT 15-DEC-1998 (Rel. 37, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Type I ribosome-inactivating protein trichoanguina precursor (rRNA
N-glycosidase) (EC 3.2.2.22) (Rip) (Trichoanguin).
GN TCA.
OS Trichosanthes angulina (Snake gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
NCBI_TaxID=50544;
RN NCBI_TaxID=50544;
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC STRAIN=cv. Angulina; TISSUE=Seed;
RX MEDLINE=99132006; PubMed=9931318;
RA Chow L.-P., Chou M.-H., Ho C.-Y., Chuang C.-C., Pan F.-M., Wu S.-H.,
Lin J.-Y.;
RT "Purification, characterization and molecular cloning of trichoanguin,
a novel type I ribosome-inactivating protein from the seeds of
Trichosanthes angulina.";
RL Biochem. J. 338:211-219(1999).
RN [2].
RP SEQUENCE OF 20-264.
RC TISSUE=Seed;
RA Chow L.-P., Kamo M., Lin J.-Y., Wang S.-H., Ueno Y., Tsugita A.;
RT "Amino acid sequence of trichoanguina, a ribosomal-inactivating
protein from Trichosanthes angulina seeds.";
RL J. Biomed. Sci. 3:178-186(1996).
CC -1- FUNCTION: INHIBITS PROTEIN SYNTHESIS BY DEPURINATING 28S RRNA IN
RIBOSOMES.
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: AF05086; AAD02686.1; -.
DR HSSP: P33185; 1BRY.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.1.
DR PRINTS: PR00396; SHIGARICIN.
DR PROSITE: PS00275; SHIGA_RICIN; FALSE_NEG.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
KW Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 264
FT PROPEP 265 294
FT ACT_SITE 177 177
FT ACT_SITE 180 180
FT CARBOHYD 70
FT CARBOHYD 220 220
FT CONFLICT 51 51
FT CONFLICT 65 65
FT CONFLICT 84 84
FT CONFLICT 152 152

TYPE I RIBOSOME-INACTIVATING PROTEIN
TRICHOANGUINA.
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (PROBABLE).
N-LINKED (GLCNAC. . .) (PROBABLE).
C -> Y (IN REF. 2).
W -> R (IN REF. 2).
N -> D (IN REF. 2).
A -> S (IN REF. 2).

FT CONFLICT 174 174 C -> S (IN REF. 2).
 SQ SEQUENCE 294 AA; 32234 MM; DA4FB8/CE3290994 CRC64;
 Query Match 21.9%; Score 312.5; DB 1; Length 294;
 Best Local Similarity 33.2%; Pred. No. 9.4e-21;
 Matches 83; Conservative 48; Mismatches 106; Indels 13; Gaps 6;

QY 4 TGDYFFETLLRDYSSGSENEIPILROSTIPVSDAQRFLVVELTNGDSTIAADY 63
 | | | | | : : : : : | | | : : : : : | | : : : : :
 DB 29 TKRYSSEFTQLRDALPTGTCVCGILPST---ASGSGFRFENLTNNDETVAVANV 85
 | | | | | : : : : : | | | : : : : : | | : : : : :
 QY 64 TNLVVAVGQDGYFLRDPAGAEHILFTGTRSSLPFGSYTLDERAG-HRQPIIG 122
 | | | | | : : : : : | | | : : : : : | | : : : : :
 DB 86 TNNVIVARADAVSYFEDTPAEAKLIFACTKVLPGSGNADKLOSVMGRQDMIELG 145
 | | | | | : : : : : | | | : : : : : | | : : : : :
 QY 123 IEEILQSVALKRPGCGSTRQAARSLIILQMISEARFNFIFRARQYINSGESFLPDWY 182
 | | | | | : : : : : | | | : : : : : | | : : : : :
 DB 146 IPALSSAITNMV--YDQSTAAALVLIQCTAEAAAYKIEQVSSHTSS--NEYPQA 201
 | | | | | : : : : : | | | : : : : : | | : : : : :
 QY 183 MLEETSWGQOSTQVQ--QSTDGVNPNPRLGISTGNFVLSNWD--VIALAIMEV 237
 | | | | | : : : : : | | | : : : : : | | : : : : :
 DB 202 VISLENNKMGALSKOIDIANRTHGOFENPELVNPDGTFVNTSAGVYKGNIKILLY 261
 | | | | | : : : : : | | | : : : : : | | : : : : :
 QY 238 CRDRSSSDY 247
 | | : : : : :
 DB 262 KAVGSEYDI 271
 | | : : : : :

RESULT 13
 RIFG_GELMU STANDARD: PRT: 316 AA.
 AC P33186; 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribosome-inactivating protein gelonin precursor (rRNA N-glycosidase)
 GN GEL.
 OS Gelonium multiflorum (Euphorbiaceae Himalaya).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Euphorbiaceae; Gelonium.
 OX NCBI_TaxID=3979;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94085781; PubMed=7916721.
 RA Nolan P.A., Garrison D.A., Belter M.;
 RT Cloning and expression of a gene encoding gelonin, a ribosome-
 RT inactivating protein from Gelonium multiflorum.";
 RL Gene 134:223-227(1993).
 RN [2]
 RP SEQUENCE OF 47-93.
 RC TISSUE=Seed;
 RX MEDLINE=89326691; PubMed=2753596;
 RA Montecuchi P.C., Lazzarini A.M., Barbieri L., Stipe F., Soria M.,
 RT Lappi D.;
 RT "N-terminal sequence of some ribosome-inactivating proteins.";
 RL Int. J. Pept. Protein Res. 33:263-267(1989).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 RX MEDLINE=9533189; PubMed=7608981.
 RA Hosur M.V., Nair B., Satyanurthy P., Misquith S., Suroli A.,
 RT Kanan K.K.;
 RT "X-ray structure of gelonin at 1.8-A resolution.";
 RL J. Mol. Biol. 250:368-380(1995).
 CC -1- CATALYTIC ACTIVITY: Endohydrolase of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 CC TYPE 1 RIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J1243; AA16312.1; -
 DR PIR: S16489; S16489.
 DR HSP: P09989; IMRJ.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP: 1.
 DR PRINTS: PS00396; SHIGARICIN.
 DR PROSITE: PS00275; SHIGARICIN.
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;
 KW Glycoprotein.
 FT SIGNAL 1 26
 FT PROPEP 27 46
 FT CHAIN 47 297
 FT PROPEP 298 316
 FT DISULFID 90 96
 FT CARBOHYD 235 235
 FT ACT_SITE 212 212
 FT CONFLICT 90 90
 FT CONFLICT 93 93
 SQ SEQUENCE 316 AA; 35418 MM; 12523E3E710901B85 CRC64;
 Query Match 21.8%; Score 311; DB 1; Length 316;
 Best Local Similarity 35.7%; Pred. No. 1.4e-20;
 Matches 86; Conservative 38; Mismatches 109; Indels 8; Gaps 4;

QY 8 YFFETLLRDYSSGSENEIPILROSTIPVSDAQRFLVVELTNGDSTIAADYTNLY 67
 | | | | | : : : : : | | | : : : : : | | : : : : :
 DB 63 YVFNELNEKVKIKRPGCNHGIPLRKNC--DQKCFVVALSDNQSLAEIADVTSYV 120
 | | | | | : : : : : | | | : : : : : | | : : : : :
 QY 68 VVAVGQDGYFLRDPAGAEHILFTGTRSSLPFGSYTLDERAGHRQPIIGIEEL 126
 | | | | | : : : : : | | | : : : : : | | : : : : :
 DB 121 VVGQVARNRSYFFKADPAAYEGIFKNTIKTRLHFGSYPSLEGKAYRETDLGIEPLR 180
 | | | | | : : : : : | | | : : : : : | | : : : : :
 QY 127 --IQSVSALRYGCGSTRQAARSLIILQMISEARFNFIFRARQYINSGESFLPDWYL 184
 | | | | | : : : : : | | | : : : : : | | : : : : :
 DB 181 IGKKIKDEADADNVPFEIASLSLVVQWSEARFTEIENGIRN--NFGQRIAPANNIT 238
 | | | | | : : : : : | | | : : : : : | | : : : : :
 QY 185 ELETSGQOSTQVQ--QSTDGVNPNPRLGISTGNFVLSNWD--VIALAIMEV 243
 | | | | | : : : : : | | | : : : : : | | : : : : :
 DB 239 SLENNKMGALSKOIDIANRTHGOFENPELVNPDGTFVNTSAGVYKGNIKILLY 298
 | | | | | : : : : : | | | : : : : : | | : : : : :
 QY 244 S 244
 | : : : : :
 DB 299 S 299
 | : : : : :

RESULT 14
 RIFL_CUCFI STANDARD: PRT: 286 AA.
 ID R1PL_CUCFI
 AC O9FRX4; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative ribosome-inactivating protein precursor (rRNA N-glycosidase)
 DE (EC 3.2.2.22).
 OS Cucumis figareli.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID=131071;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada T., Ohki S.T., Osaki T.;
 RT Cloning and analysis of a cDNA coding a putative ribosome-
 RT inactivating protein from Cucumis figareli.";
 RL Plant Biotechnol 17:337-340(2000).
 CC -1- CATALYTIC ACTIVITY: Endohydrolase of the N-glycosidic bond at one

CC specific adenosine on the 28S rRNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 CC TYPE 1 RIP SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB045560; BAB19677.1; -
 DR HSSP: P16094; 1AHC.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP.1.
 DR PRINTS: PR00396; SHIGARICIN.
 DR PROSITE: PS00275; SHIGA_RICIN.1.
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
 KW Toxin; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 286
 FT ACT_SITE 185 185
 FT CARBOHYD 103 103
 FT CARBOHYD 110 110
 FT CARBOHYD 252 252
 SQ SEQUENCE 286 AA; 31771 MW; 4EPD4966E04DA41 CRC64;
 Query Match 21.6%; Score 308.5; DB 1; Length 286;
 Best local Similarity 33.2%; Pred. No. 2.1e-20;
 Matches 78; Conservative 52; Mismatches 92; Indels 13; Gaps 7;
 QY 8 YFFRLLRDYVSSGFSNEIFLRQSTIPYSDAQRFLVELFNGGSIADINNG 67
 DB 40 YSFRTSMRALPAGNDIYVLPSP--ISGSRITLMQSNTEGNTINADYVNY 96
 QY 68 VVAYOAGDOSYFLRDA-PDGAERHLFTGTRSSLPFTGTYDLERYA-GHRDQIPLGTEE 125
 DB 97 IMGYLVNGTSTYFNETDQALASKVFQGTSTKSTLPYSGNVOKRQSVARERPSDPLGMA 156
 QY 126 LIOSVALRYPGSTRQAARSLIILQIMISBAARFNIFWRAKOYINS-GESEFLPDWML 184
 DB 157 LDSAISTLVY--YDSRSAPIAFVLICITAEAAAYKYI--EKQIIDRISVSKVDLAAI 211
 QY 185 ELFTSMGQOSTOVQ--QSTDDGVENPFRLGISTGNFVLTSSNVADVASLALMLFV 237
 DB 212 SLENEMSLSKQIQAKSNKGQFQTPVKIINDKGLTEVTNNSLVVTKNTMLL 266
 RESULT 15
 RIPL_BRYDI STANDARD: PRT; 290 AA.
 AC P3185; Q9S819.
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)
 DE (EC 3.2.2.22) (B01).
 OS Bryonia dioica (Red bryonia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eucrosts I; Cucurbitales; Cucurbitaceae; Bryonia.
 OC NCBI_TaxID=3652;
 RN [1]
 RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RC TISSUE=Leaf;
 RX MEDLINE=97228081; PubMed=9115985;
 RA Gwialak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,
 RA Siegel C.B.,
 RT "Molecular, biological, and preliminary structural analysis of
 RT recombinant bryodin I, a ribosome-inactivating protein from the plant
 RT Bryonia dioica".
 RL Biochemistry 36:3095-3103(1997).

RN [2]
 RP SEQUENCE FROM N.A.
 RA Siegel C.B.,
 RT "Cloning and expression of a gene encoding bryodin I from Bryonia
 RT dioica".
 RL Patent number US5541110, 30-JUL-1996.
 RP SEQUENCE OF 24-66.
 RC TISSUE=seed;
 RX MEDLINE=89326691; PubMed=2753596;
 RA Montecucchi P.C., Lazzerini A.M., Barbieri L., Stirpe F., Soria M.,
 RA Lappi D.,
 RT "N-terminal sequence of some ribosome-inactivating proteins".
 RL Int. J. Pept. Protein Res. 33:263-267(1989).
 RP SEQUENCE OF 24-43.
 RC TISSUE=ROOT;
 RX MEDLINE=95151812; PubMed=7849072;
 RA Siegel C.B., Gwialak S.L., Chace D., Wolff E.A., Mixan B.,
 RA Marguardt H.,
 RT "Characterization of ribosome-inactivating proteins isolated from
 RT Bryonia dioica and their utility as carcinoma-reactive
 RT Immunokonjugates".
 RL Bioconjug. Chem. 5:423-429(1994).
 CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS.
 CC -1- CATALYTIC ACTIVITY: Endohydrolase of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -1- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO
 CC PRODUCE A SHORTER PROTEIN.
 CC -1- BIOCHEMISTRY: Especially useful as immunotoxin for
 CC pharmacological applications as it has low toxicity in rats and
 CC mice but is potent once inside target cells.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 CC TYPE 1 RIP SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: J24020; -; NOT ANNOTATED_CDS.
 DR PIR: S16491; S16491.
 DR PDB: 1BRV; 04-MAR-98.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP.1.
 DR PRINTS: PR00396; SHIGARICIN.
 DR PROSITE: PS00275; SHIGA_RICIN.1.
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
 KW 3D-structure; Multigene family; Glycoprotein; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 290
 FT PROPEP 271 270
 FT ACT_SITE 183 183
 FT ACT_SITE 212 212
 FT CARBOHYD 214 214
 FT CARBOHYD 250 250
 FT MTAAEN 212 212
 FT CONFLICT 61 65
 SQ SEQUENCE 290 AA; 31788 MW; E96CDDC031AA2DB CRC64;
 Query Match 20.9%; Score 297.5; DB 1; Length 290;
 Best local Similarity 32.9%; Pred. No. 2.1e-19;
 Matches 82; Conservative 54; Mismatches 88; Indels 25; Gaps 9;
 QY 5 GDEYFR-----FTLLRDYVSSGFSNEIFLRQSTIPYSDAQRFLVELFNGGSIADINNG 53
 DB 23 GDVSPRLSGATTTGSGVFIKMLREALPYERKYNIPILRSS--ISGSGRYTLHLHTNYA 79
 QY 54 GDSITAIIDVTLVYVAYOAGDOSYFLRDA-PDGAERHLFTGTRSS-LPFTGTYDLER 111

Thu Mar 27 07:07:03 2003

us-09-627-165e-16.rsp

Page 13

[illegible]

Search completed: March 18, 2003, 08:25:46
Job time : 15.6181 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:23:18 ; Search time 37.1533 Seconds
(without alignments)
1547.297 Million cell updates/sec

Title: US-09-627-165E-16
Perfect score: 1425
Sequence: 1 HQTGDEYFRITLLRDYVS.....NSGAVDDVTCTASEPTVRIV 279

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	1417	99.4	565	10 Q8W243	Q8W243 viscum albu
2	1264.5	86.7	531	10 Q8RXH6	Q8RXH6 viscum albu
3	1187	83.3	249	10 Q8RXH7	Q8RXH7 viscum albu
4	545	38.2	580	10 Q94BW3	Q94BW3 cinamomum
5	536	37.6	549	10 Q9FV22	Q9FV22 cinamomum
6	535	37.5	581	10 Q94BW4	Q94BW4 cinamomum
7	481	33.8	528	10 Q06076	Q06076 abrus precia
8	480	33.7	252	10 Q38761	Q38761 abrus precia
9	474	33.2	251	10 Q96237	Q96237 abrus precia
10	473	33.2	251	10 Q96235	Q96235 abrus precia
11	467	33.1	252	10 Q38760	Q38760 abrus precia
12	458	30.1	592	10 Q8W2E7	Q8W2E7 liris hollan
13	428.5	30.0	570	10 Q41358	Q41358 sambucus ni
14	427	29.9	573	10 Q8W2E8	Q8W2E8 liris hollan
15					
16					

17	425	29.8	564	10 Q9AVR2	Q9AVR2 sambucus eb
18	422.5	29.6	541	10 Q41174	Q41174 rictinus com
19	422.5	29.6	565	10 Q04071	Q04071 sambucus ni
20	421	29.5	570	10 Q22415	Q22415 sambucus ni
21	416	29.2	547	10 Q9M6E9	Q9M6E9 abrus precia
22	403	28.3	563	10 Q04367	Q04367 sambucus ni
23	401.5	28.2	604	10 Q9M654	Q9M654 polyonatum
24	390	27.4	563	10 Q94552	Q94552 sambucus ni
25	387.5	27.2	569	10 P93543	P93543 sambucus ni
26	383	26.9	566	10 Q04072	Q04072 sambucus ni
27	356.5	25.0	603	10 Q9M653	Q9M653 polyonatum
28	338.5	23.8	286	10 Q9FUV7	Q9FUV7 memordica c
29	337	23.6	289	10 Q94KE4	Q94KE4 trichosanthe
30	333.5	23.4	286	10 Q41257	Q41257 memordica c
31	332.5	23.3	254	10 Q9F5H2	Q9F5H2 memordica c
32	329	23.1	289	10 Q41216	Q41216 memordica c
33	327.5	23.0	300	10 Q04357	Q04357 liris hollan
34	324	22.7	247	10 Q9LRE3	Q9LRE3 trichosanthe
35	319.5	22.4	300	10 Q04356	Q04356 liris hollan
36	303.5	21.3	270	10 Q41611	Q41611 trichosanthe
37	303	21.3	258	10 Q04358	Q04358 liris hollan
38	302.5	21.2	258	10 Q9S9E4	Q9S9E4 gelonium mu
39	291.5	20.5	278	10 Q00980	Q00980 lufia cylin
40	290.5	20.4	293	10 Q8S452	Q8S452 jatrophe cu
41	289.5	20.3	293	10 Q8VYU0	Q8VYU0 jatrophe cu
42	263.5	18.5	305	10 Q8W4U4	Q8W4U4 bougainvill
43	227	15.9	365	10 Q8RYE9	Q8RYE9 gynostemma
44	226.5	15.9	382	10 Q8SA43	Q8SA43 abrus precia
45	212	14.9	302	10 P93077	P93077 clerodendru

ALIGNMENTS

```
RESULT 1
ID Q8W243 PRELIMINARY: PRT: 565 AA.
AC Q8W243:
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE VCA precursor.
OS Viscum album subsp. coloratum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_TaxID=159976;
RN [1]
RP SEQUENCE FROM N.A.
RA Park W.-B., Lyu S.;
RT *Cloning of Viscum album subsp. coloratum (Korean mistletoe).";
RL Biochem. Biophys. Res. Commun. 0:0-0(2002).
DR EMBL: AF369961; AAL40417.1; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR000772; Ricin_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00652; Ricin_B_lectin; 6.
DR Pfam: PF00161; RIP; 1.
DR PRINTS: PR00396; SHIGARICIN.
DR SMART: SM00458; RICIN; 2.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS50231; RICIN_B_LECTIN; 2.
KW Signal.
FT SIGNAL. 1 22 POTENTIAL.
FT CHAIN. 23 273 VCA ALPHA CHAIN.
FT CHAIN. 309 565 VCA BETA CHAIN.
SQ SEQUENCE 565 AA; 62401 MW; 991E3994D0A005F11 CRC64;

Query Match 99.4%; Score 1417; DB 10; Length 565;
Best Local Similarity 99.6%; Pred. No. 4.9e-121;
Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 HQTGDEYFRITLLRDYVSFGSFSNEIPLRQSTIPVSDAQRFLVELTNGGDSITAA 60

```

Db 43 HQTGDEYFRFTLLRDYVSSGFSNFIPLLRSTIPVSDAORFVLELTNOGDSITAA 102
    |||
Qy 61 IDVTNLYVAVAGDQSYFLDAPDGAERHLFTGTRSSLPTGSGYTDLERYAGHRQIP 120
    |||
Db 103 IDVTNLYVAVAGDQSYFLDAPDGAERHLFTGTRSSLPTGSGYTDLERYAGHRQIP 162
    |||
Qy 121 LGTEELIOSVALRPGSGTAAORSLITLQMSAARFNIPFARQYINSGESFLPD 180
    |||
Db 163 LGTEELIOSVALRPGSGTAAORSLITLQMSAARFNIPFARQYINSGESFLPD 222
    |||
Qy 181 MYMLELTSWGOOSTOVQOSTDGVFNPFRLGISTGNEVTLISNVDYASIALMLFVCRD 240
    |||
Db 223 MYMLELTSWGOOSTOVQOSTDGVFNPFRLGISTGNEVTLISNVDYASIALMLFVCRD 282
    |||
Qy 241 RPSSSDRYWPLVIRPVLENSGAVDVTCTASEPTVRIY 279
    |||
Db 283 RPSSSDRYWPLVIRPVLENSGAVDVTCTASEPTVRIY 321
    |||

```

RESULT 2

```

ID 08RXH6 PRELIMINARY: PRT: 531 AA.
AC 08RXH6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Lectin chain A isoform 1 (Fragment).
OS Viscum album (European mistletoe).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_TaxID=3972;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Paramasivam M., Mista V., Sriinivasan A., Singh T.P.;
RT "Viscum album (Indian) mRNA for Mistletoe lectin chain A isoform 1 and
    chain B."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY081149; AAL87006.1; -.
FT NON_TER 1 249 LECTIN CHAIN A ISOFORM 1.
FT CHAIN 269 531 LECTIN B CHAIN.
FT CHAIN 531 531
SQ SEQUENCE 531 AA; 58802 MW; 18244BEEFE35422 CRC64;

```

Query Match 88.7%; Score 1264.5; DB 10; Length 531;
 Best Local Similarity 88.2%; Pred. No. 4e-107;
 Matches 246; Conservative 15; Mismatches 13; Indels 5; Gaps 1;

```

Qy 1 HQTGDEYFRFTLLRDYVSSGFSNFIPLLRSTIPVSDAORFVLELTNOGDSITAA 60
    |||
Db 10 HQTGDEYFRFTLLRDYVSSGFSNFIPLLRSTIPVSDAORFVLELTNOGDSITAA 69
    |||
Qy 61 IDVTNLYVAVAGDQSYFLDAPDGAERHLFTGTRSSLPTGSGYTDLERYAGHRQIP 120
    |||
Db 70 IDVTNLYVAVAGDQSYFLDAPDGAERHLFTGTRSSLPTGSGYTDLERYAGHRQIP 129
    |||
Qy 121 LGTEELIOSVALRPGSGTAAORSLITLQMSAARFNIPFARQYINSGESFLPD 180
    |||
Db 130 LGTEELIOSVALRPGSGTAAORSLITLQMSAARFNIPFARQYINSGESFLPD 189
    |||
Qy 181 MYMLELTSWGOOSTOVQOSTDGVFNPFRLGISTGNEVTLISNVDYASIALMLFVCRD 240
    |||
Db 190 MYMLELTSWGOOSTOVQOSTDGVFNPFRLGISTGNEVTLISNVDYASIALMLFVCRD 249
    |||
Qy 241 RPSSSDRYWPLVIRPVLENSGAVDVTCTASEPTVRIY 279
    |||
Db 250 RPSSSDRYWPLVIRPVLENSGAVDVTCTASEPTVRIY 283
    |||

```

RESULT 3

08RXH7

```

ID 08RXH7 PRELIMINARY: PRT: 249 AA.
AC 08RXH7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Lectin chain A isoform 2 (Fragment).
OS Viscum album (European mistletoe).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Viscaceae; Viscum.
OX NCBI_TaxID=3972;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Paramasivam M., Sriinivasan A., Singh T.P.;
RT "Viscum Album (Indian) mRNA for Mistletoe Lectin Chain A, isoform 2."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY081148; AAL87005.1; -.
FT NON_TER 1 249
FT NON_TER 1 249
SQ SEQUENCE 249 AA; 27944 MW; 89FAFB78309A83B3 CRC64;

```

Query Match 83.3%; Score 1187; DB 10; Length 249;
 Best Local Similarity 97.9%; Pred. No. 1.7e-100;
 Matches 233; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

Qy 1 HQTGDEYFRFTLLRDYVSSGFSNFIPLLRSTIPVSDAORFVLELTNOGDSITAA 60
    |||
Db 10 HQTGDEYFRFTLLRDYVSSGFSNFIPLLRSTIPVSDAORFVLELTNOGDSITAA 69
    |||
Qy 61 IDVTNLYVAVAGDQSYFLDAPDGAERHLFTGTRSSLPTGSGYTDLERYAGHRQIP 120
    |||
Db 70 IDVTNLYVAVAGDQSYFLDAPDGAERHLFTGTRSSLPTGSGYTDLERYAGHRQIP 129
    |||
Qy 121 LGTEELIOSVALRPGSGTAAORSLITLQMSAARFNIPFARQYINSGESFLPD 180
    |||
Db 130 LGTEELIOSVALRPGSGTAAORSLITLQMSAARFNIPFARQYINSGESFLPD 189
    |||
Qy 181 MYMLELTSWGOOSTOVQOSTDGVFNPFRLGISTGNEVTLISNVDYASIALMLFVCRD 238
    |||
Db 190 MYMLELTSWGOOSTOVQOSTDGVFNPFRLGISTGNEVTLISNVDYASIALMLFVCRD 247
    |||

```

RESULT 4

094BW3

```

ID 094BW3 PRELIMINARY: PRT: 580 AA.
AC 094BW3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE rRNA - glycosylase precursor (EC 3.2.2.22).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
    genes encoding cinnamomum proteins and study of their expression
    patterns."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDONUCLEOLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
    SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
DR EMBL: AY039803; AAK82460.1; -.
DR InterPro: IPR000772; RICHN_B_lectin.
DR InterPro: IPR001574; RIP.
DR Pfam: PF00161; RIP.1;
DR Pfam: PF00161; RIP.1;
DR PROSITE: PS0231; RICHN_B_LECTIN; 2.
DR Hydrolase; Signal; Toxin.
KW SIGNAL 1 32 POTENTIAL.
FT SIGNAL 1 32

```

[illegible]

RESULT	ID	SEQUENCE	PRELIMINARY	PROT.	549 AA.
09FV22	09FV22				
AC	09FV22				
DT	01-MAR-2001	(TrEMBLrel. 16, Created)			
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)			
DT	01-MAR-2002	(TrEMBLrel. 20, Last annotation update)			
DE	tRNA - glycosylase (EC 3.2.2.22) (Fragment).				
OS	Cinnamomum camphora (Camphor tree).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.				
OX	NCBI_TaxID=13429;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Xie L., Liu W.-Y., Wang E.-D.;				
RT	"Molecular cloning of cinnamomyl A-, B-chain and the expression,				
RT	purification, characterization and mutagenesis of the A-Chain.";				
RL	Submitted (SEP-2000) to the EMBL/Genbank/DBD databases.				
CC	-1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE				
CC	SPECIFIC ADENOSINE ON THE 28S RRNA.				
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.				
CC	EMBL: AE259548; AAF68978.2; "				
DR	HSSP: P02879; ZAAI.				
DR	InterPro: IPR0000772; Ricin_B_lectin.				
DR	InterPro: IPR001574; RIP.				
DR	Pfam: PF006652; Ricin_B_lectin; 5.				
DR	Pfam: PF00161; RIP; 1.				
DR	PRINTS: PRO0396; SHIGARICIN.				
DR	SMART: SM00458; Ricin_2.				
DR	PROSITE: PS50231; Ricin_B_LECTIN; 2.				
SW	Hydrolase; Toxin.				
FT	NON_TER				
FT	1				
SO	SEQUENCE				
SO	549 AA; 60648 MW; 02607FE607CA44B0 CRC64;				
Query Match					
Best Local Similarity	46.7%;				
Matches	136; Conservative 44; Mismatches 95; Indels 16; Gaps 10;				
OY	2 QTTGDEYERFTLLRDYVSSGSNFIPLRLQ-STIPVSAQRFVVELITNOGDS-TYA 59				
DB	10 KAKTSYIOFIDLRAGLASGEHPHGIPIVRRSTVP-DSKRFILVELSNMAADSPVTL 67				
OY	60 AIDVTVVYVQAGDQSYFLR-DAPDGAERHLEFTCTRRSLPFTSYTDERVAG-HRD 117				

Db 68 ADVYNAVYARTRGSGFLEBNPPALENLPTKRYTFPPSSGYTDLGAVGERE 127

Qy 118 QIPDIETELTQSALRYRG-GRPAOABSLITLIMISEAPRNPITPRAQYINSES 176

Db 128 ELLDNQDLENLASHALMISNLKQBPALRSRLIYIOMVAEVRFEIYKOSISRBEM 187

Qy 177 FLPPDMLELETSNGOOSTQVOQSTD-GYNNPFRLGISTGNYTSLNVRD-VYASLAIM 234

Db 188 FRPDPAHLSLENNKMSALNAVQSGSGSSVEPLRSISKNPYVQSVSDRYISLAIM 247

Qy 235 LEVGR--DRPSSDVRKPLVTRPYLEN-----SGADVDTCTASEPTVRI 278

Db 248 LPTCTSDTRASDQFLIDHMLMLRPLVDVAEVAITDMDNDTCCAPEPTVRI 298

RESULT 6	094BM4	PRELIMINARY:	PRT:	580 AA.
ID	094BM4			
AC	094BM4			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	rRNA - glycosylase precursor (EC 3.2.2.22).			
OS	Cinnamomum camphora (Camphor tree).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.			
OX	NCBI_TaxID=134429;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Yang O., Gong Z.Z., Liu W.Y.;			
RT	"Molecular cloning of three type 2 rtp (ribosome-inactivating protein			
RT	genes encoding cinnamomum proteins and study of their expression			
RT	patterns."			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
CC	-I- CATALYTIC ACTIVITY: ENDORHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONHR			
CC	SPECIFIC ADENOSINE BELONGS TO THE 28S RRNA.			
CC	-I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.			
DR	EMBL; AY039802; AAK8459.1; "			
DR	InterPro; IPR000772; R1c1n_B_lectin.			
DR	InterPro; IPR001574; R1P.			
DR	Pfam; PF00652; R1c1n_B_lectin; 6.			
DR	Pfam; PF00161; R1P; 1.			
DR	PROSITE; PSS0231; R1CIN_B_LECTIN; 2.			
DR	Hydrolase; Signal; Toxin.			
FT	SIGNAL	1	32	POTENTIAL.
FT	CHAIN	33	580	TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT				CINNAMOMIN II.
QO	SEQUENCE	580 AA;	64265 MW;	3784289ECCECCGFF CRC64;

Query Match	37.6%	Score 536	Dh 10	Length 580
Best Local Similarity	47.1%	Ped. No. 2	Ie 40	
Matches 136	Conservative 42	Mismatches 95	Indels 16	Gaps 10
QY	4	TGGEFFRFTLLRRYVSSGSFNSNEIPILRO--STPVSDAOGRFLVLTMOGDS--ITAAI	61	
Dh	44	TKSTSTGFIETLRQAQLGAEHPHG:VWRKDSTVP--DSKRFLIVELSMWADSPVALAV	101	
QY	62	DVTNLYVAAYAGQSGSYELR-DAPDGADEHHFTGTSRSLPFGSGTDLDERAGH-RDOI	119	
Dh	102	DVTNLYVAAYRTGSGSFLEDRDPPRAIENLLPTKRYTTPFGSGTDLDERAGAGLEEI	161	
QY	120	PGIEELTQSSALRYPG--GSTRPAQRSLITLIMIGEARFNPIMWARYOYINGSGFL	178	
Dh	162	LGMDFLENAISALMTSNLMOQALARSLIVYIQWAEAVRFRIEYRRESEITPAEEFR	221	
QY	179	PKMYMLETETSMGQSGTQVOGSTD--GVNNPFRIGISTGFIYLSVNRV-VIASLALIMF	236	
Dh	222	PPRAMLSLFNKWSALSMAYQSGNVGSFSPVELRISLNNKYVGVSGDSRYISGLALIMF	281	
QY	237	VCR--DRPSSDVAKYKFLVRYPLEK-----SGAVDDVDTQTAASEPVRI	278	
Dh	282	ICRSSDRTSSDQFIIDLMLIRPLIVYADVATDADNDNDTCADPEEPIYRI	330	

RESULT 7

Q94BW5 PRELIMINARY: PRT; 581 AA.

AC Q94BW5; PRELIMINARY: PRT; 581 AA.
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE rRNA -glycosidase precursor (EC 3.2.2.22).
 OS Cinnamomum camphora (Camphor tree).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: Laurales: Lauraceae: Cinnamomum.
 NCBI_TaxID=13429;

RP SEQUENCE FROM N.A.

KA Yang Q., Gong Z.Z., Liu W.Y.;

RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein) genes encoding cinnamomin proteins and study of their expression patterns."

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

CC -1 CATALYTIC ACTIVITY: ENDOTHEROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE SPECIFIC ADENOSINE ON THE 28S RNA.

CC -1 SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.

DR EMBL: AY039801; AAK82458.1; "

DR InterPro: IPR000772; R1CIN_B_Lectin.

DR Pfam: PF00652; R1CIN_B_Lectin; 5.

DR PROSITE: PS00231; R1CIN_B_Lectin; 2.

DR HydroLase: Signal; Toxin.

FT CHAIN 1 32 POTENTIAL.

FT SIGNAL 1 32 TYPE 2 RIBOSOME-INACTIVATING PROTEIN.

SQ SEQUENCE 581 AA; 64215 MW; 6E8F5FBFB3D196 CRC64;

Query Match 37.5%; Score 535; DB 10; Length 581;

Best Local Similarity 47.1%; Pred. No. 2.5e-40;

Matches 136; Conservative 43; Mismatches 94; Indels 16; Gaps 10;

QY 4 TGDEFFRITLDRVSSGSFSENEIPLRO-STIPVDAORFVELTNOGDS-ITAI 61
 DB 4 TKTSTQPIELRAQLASGEPRGIPVREKSTVP--DSKRFIVELSNMADSVYTLAV 101
 QY 62 DVTNLYVAYOAGQSYFLR-DAPDCAERHFTGTSSLPFGSTTDLEKAG-HRQI 119
 DB 102 DVTNLYVAYRTGSQSEFLEDPDALENLDPDKRTPEFGSTTDLEKAGGERREI 161
 QY 120 PLGIBELIQSVSALRYPG-GSTRAQARSLILLIOMISEARFNPFIWRAQYINSGESFL 178
 DB 162 ILGMDPLENAISALWISNLNOQARALARSLIVVQVAVAFPIEYRVRGSSISRAEMFR 221
 QY 179 PDWYMLELETISWGQSTOVQSTD-GVFNPFRLGISTGNFVTLNVRD-VIASLAIMLF 236
 DB 222 FDPDAWLSLENKWSALSNVQSGVGFSPVELRISINRKYVYGVSDRVSIGLAIMLF 281
 QY 237 VCR--DRPSSDYRWPIVIRPVLEN-----SGAVDVTCTASEPTVRI 278
 DB 282 ICRSDIRASSDQFDIMIMIRPIVLDVAEVATDADNDCTADEPPIVRI 330

RESULT 8

Q06076 PRELIMINARY: PRT; 528 AA.

AC Q06076; PRELIMINARY: PRT; 528 AA.
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
 OC Eucosids I: Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 NCBI_TaxID=3816;

RESULT 9

Q38761 PRELIMINARY: PRT; 252 AA.

AC Q38761; PRELIMINARY: PRT; 252 AA.
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
 OC Eucosids I: Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 NCBI_TaxID=3816;

RP SEQUENCE FROM N.A.

KA Yang Q., Gong Z.Z., Liu W.Y.;

RT "Primary structure of three distinct isoforms determined by cDNA sequencing: conservation and significance."

RL J. Mol. Biol. 229:263-267(1993).

CC -1 CATALYTIC ACTIVITY: ENDOTHEROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE SPECIFIC ADENOSINE ON THE 28S RNA.

CC -1 SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.

DR EMBL: M98346; AAA32626.1; "

DR HSP: P11440; IABR.

DR InterPro: IPR000772; R1CIN_B_Lectin.

DR Pfam: PF00652; R1CIN_B_Lectin; 6.

DR SMART: SM00458; R1CIN; 2.

DR PROSITE: PS00231; R1CIN_B_Lectin; 2.

DR PROSITE: PS00275; SHIG_R1CIN; UNKNOWN_1.

DR HydroLase: Toxin.

FT NON_TER 1 1

FT NON_TER 1 1

SQ SEQUENCE 528 AA; 58870 MW; 62ED42FB8FE60F8 CRC64;

Query Match 33.8%; Score 481; DB 10; Length 528;

Best Local Similarity 43.7%; Pred. No. 1.9e-35;

Matches 121; Conservative 41; Mismatches 103; Indels 12; Gaps 7;

QY 4 TGDEFFRITLDRVSSGSFSENEIPLRO-STIPVDAORFVELTNOGDSITAI 63
 DB 13 TSQSYKQFELARQL-TGGLIHIDIPVLPDP-TVEERNRYIVELSERESTLENGID 70
 QY 64 TNLVYVAYOAGQSYFLRDAPDAERHFTGTSSLPFGSTYDLEKAG-HRQI 122
 DB 71 TNLVYVAYAGQSYFLRDAPDAERHFTGTSSLPFGSTYDLEKAG-HRQI 130
 QY 123 IEEIIOVSALRYPGSTRAQARSLILLIOMISEARFNPFIWRAQYINSGESFL 182
 DB 131 LQALTALIFLNSGASNDKARFLVYIOMSEARFNPFIWRAQYINSGESFL 190
 QY 183 MLELETISWGQSTOVQSTDGVF--NPFRLGISTGNFVTLNVRD-VIASLAIMLF 241
 DB 191 MLELETISWGQSTOVQSTDGVF--NPFRLGISTGNFVTLNVRD-VIASLAIMLF 249
 QY 242 PSSSDYRWPIVIRPVLEN--SGAVDVTCTASEPTVRI 278
 DB 250 PNMAG--PLIRISIVES---KICSSRYEPTVRI 279

RESULT 9

Q38761 PRELIMINARY: PRT; 252 AA.

AC Q38761; PRELIMINARY: PRT; 252 AA.
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
 OC Eucosids I: Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 NCBI_TaxID=3816;

RA Chh-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
 "Cloning and expression of three abrin A-chains and their mutants
 derived by site-specific mutagenesis in *Escherichia coli*.";
 Eur. J. Biochem. 219:83-87(1994).
 RL -1- CATALYTIC ACTIVITY: ENDOPHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: X54873; CA38655.1; -
 DR EMBL: P11140; IABR.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP.
 DR PRINTS: PR00396; SHIGARICIN.
 DR PROSITE: PS00275; SHIGA_RICIN.1.
 KW Glycosidase; Hydrolyase; Toxin.
 FT NON_TER 252
 SQ SEQUENCE 252 AA; 28229 MW; 187BB84E134AEC5 CRC64;
 Query Match 33.7%; Score 480; DB 10; Length 252;
 Best Local Similarity 44.1%; Pred. No. 8.2e-36;
 Matches 109; Conservative 36; Mismatches 78; Indels 24; Gaps 5;
 QY 4 TGDEYFRFTLLRDVYSSGSFSNEIPLRQSTIPVSDAQRVLYELTNGGDSITAAIDV 63
 DB 14 TSQSYKQFIALERL-RGGLIHDIPLDPT-TLQENRRITVELSNSDTSIEVGIDV 71
 QY 64 TNLVYVAYAGDSYFLRDAPDAERHLEFGTTRSSLPFGSTDLERYAGH-RDQIPLG 122
 DB 72 TNAIVYVAYAGDSYFLRDAPDAERHLEFGTTRSSLPFGSTDLERYAGH-RDQIPLG 131
 QY 123 IEELIOSVALRPGSGSTRAGASLIILOMISEARPNPFMRARQYINSGESFLPDMY 182
 DB 132 LQALTHGISFRRSGNDNEKARTLLVIIOYAAARRYSNKNVYSIQGTAFQPDAA 190
 QY 183 MLEETSMGQSTVOQSTDGVFNPRFGISTGVFTLSNVRD-----VIASL 231
 DB 192 MISLENNMNLGRGVQESVDTFPNO-----VTLINIRNRPYVDSLSHPTAVL 240
 QY 232 AIMLFVC 238
 DB 242 ALMLFVC 248
 RESULT 10
 ID 096236 PRELIMINARY; PRT; 251 AA.
 AC 096236;
 RX MEDLINE=94139756; PubMed=8307038;
 RA Chh-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
 "Cloning and expression of three abrin A-chains and their mutants
 derived by site-specific mutagenesis in *Escherichia coli*.";
 Eur. J. Biochem. 219:83-87(1994).
 RL -1- CATALYTIC ACTIVITY: ENDOPHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: X76721; CA54139.1; -
 DR EMBL: P11140; IABR.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP.
 DR PRINTS: PR00396; SHIGARICIN.

KW Hydrolyase; Toxin.
 FT NON_TER 1
 FT NON_TER 251
 SQ SEQUENCE 251 AA; 28040 MW; D57F81B2E0ECC9 CRC64;
 Query Match 33.3%; Score 474; DB 10; Length 251;
 Best Local Similarity 43.7%; Pred. No. 2.9e-35;
 Matches 108; Conservative 36; Mismatches 79; Indels 24; Gaps 5;
 QY 4 TGDEYFRFTLLRDVYSSGSFSNEIPLRQSTIPVSDAQRVLYELTNGGDSITAAIDV 63
 DB 13 TSQSYKQFIALERL-RGGLIHDIPLDPT-TLQENRRITVELSNSDTSIEVGIDV 70
 QY 64 TNLVYVAYAGDSYFLRDAPDAERHLEFGTTRSSLPFGSTDLERYAGH-RDQIPLG 122
 DB 71 TNAIVYVAYAGDSYFLRDAPDAERHLEFGTTRSSLPFGSTDLERYAGH-RDQIPLG 130
 QY 123 IEELIOSVALRPGSGSTRAGASLIILOMISEARPNPFMRARQYINSGESFLPDMY 182
 DB 131 LQALTHGISFRRSGNDNEKARTLLVIIOYAAARRYSNKNVYSIQGTAFQPDAA 190
 QY 183 MLEETSMGQSTVOQSTDGVFNPRFGISTGVFTLSNVRD-----VIASL 231
 DB 191 MISLENNMNLGRGVQESVDTFPNO-----VTLINIRNRPYVDSLSHPTAVL 240
 QY 232 AIMLFVC 238
 DB 241 ALMLFVC 247
 RESULT 11
 ID 096237 PRELIMINARY; PRT; 251 AA.
 AC 096237;
 RX MEDLINE=94139756; PubMed=8307038;
 RA Chh-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
 "Cloning and expression of three abrin A-chains and their mutants
 derived by site-specific mutagenesis in *Escherichia coli*.";
 Eur. J. Biochem. 219:83-87(1994).
 RL -1- CATALYTIC ACTIVITY: ENDOPHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: X76722; CA54140.1; -
 DR EMBL: P11140; IABR.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP.
 DR PRINTS: PR00396; SHIGARICIN.
 KW Hydrolyase; Toxin.
 FT NON_TER 1
 FT NON_TER 251
 SQ SEQUENCE 251 AA; 28055 MW; 6F64755C3DEAB79 CRC64;
 Query Match 33.2%; Score 473; DB 10; Length 251;
 Best Local Similarity 43.7%; Pred. No. 3.6e-35;
 Matches 108; Conservative 36; Mismatches 79; Indels 24; Gaps 5;
 QY 4 TGDEYFRFTLLRDVYSSGSFSNEIPLRQSTIPVSDAQRVLYELTNGGDSITAAIDV 63
 DB 13 TSQSYKQFIALERL-RGGLIHDIPLDPT-TLQENRRITVELSNSDTSIEVGIDV 70
 QY 64 TNLVYVAYAGDSYFLRDAPDAERHLEFGTTRSSLPFGSTDLERYAGH-RDQIPLG 122

Thu Mar 27 07:07:04 2003

us-09-627-165e-16.rspt

Page 6

Db 71 TNNVAVKRACTGSGYFLRDAAPSSADYLFPTTDDHSLPFGYTGDLERMAHNSQDTPG 130
 123 IEEILQSLALRTYGGSTRAQASLILILQWISAAFNPIWRARQYINSGESFLPMY 182
 Db 131 LQALHTGISTFRRSGNDNEKARTLLVILQWAAALFRTISNRVAVSLQGTAFQDDA 190
 QY 183 MLEETSSNGGOSTVOQSTDGYFNNPFLGISTGNFTYLSNVND-----VYASL 231
 Db 191 MSLSLNNMNDNLNRGVQDSYVDTFPNO-----VTLTINIRNEPIVDSLHPYAVL 240
 QY 232 AITMLFVC 238
 Db 241 AITMLFVC 247
 RESULT 12
 096235
 ID Q96235 PRELIMINARY; PRT: 251 AA.
 AC 096235;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE rRNA -glycosidase (EC 3.2.2.22) (Fragment).
 OS Abrus precatorius (Indian jicorice) (Crab's eye).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Ochromatophyta; Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae;
 CC OC
 CC NCBI_TaxID:3816;
 RN
 RN
 RN SEQUENCE FROM N.A.
 RC
 RC TISSUE=SEED;
 RA MEDLINE=94139756; PubMed=8307038;
 RA Chih-Hung H., Lee M.C., Chen J.K., Lin J.Y.;
 RT Cloning and expression of three abrin A-chains and their mutants
 RT derived by site-specific mutagenesis in *Escherichia coli*.
 RL Eur. J. Biochem. 219:83-87(1994).
 CC
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RRNA.
 CC
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY.
 DR EMBL: X76644; CAA54092.1; -.
 DR HSSP: P1140; IABR.
 DR Interpro: IPR001574; RIP.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PR00396; SHIGARICIN.
 DR Hydrolase; Toxin.
 KW
 KW NON_TER 1
 FT NON_TER 251
 FT 251
 SQ SSOURCE 251 AA; 27997 MW; 3B60351839AEF7E CRC64;
 Query Match 32.8%; Score 467; DB 10; Length 251;
 Best Local Similarity 43.3%; Pred. No. 1,3e-14;
 Matches 107; Conservative 36; Mismatches 80; Indels 24; Gaps 5;
 QY 4 TGEYERFTTLLRDVSSGSFSNSPEILRQSTIPSAORFVVELTNGGDSITAAIDV 63
 Db 13 TQSGYKQFTEALRELT-RGGJLHDIYVLPDP-TLQERNRYITVELSNDSTSEISYGDV 70
 QY 64 TNNVAVYAGQSGSYFLRDAAPSSADYLFPTTDDHSLPFGYTGDLERMAHNSQDTPG 130
 Db 71 TNNVAVYAGQSGSYFLRDAAPSSADYLFPTTDDHSLPFGYTGDLERMAHNSQDTPG 130
 QY 123 IEEILQSLALRTYGGSTRAQASLILILQWISAAFNPIWRARQYINSGESFLPMY 182
 Db 131 LQALHTGISTFRRSGNDNEKARTLLVILQWAAALFRTISNRVAVSLQGTAFQDDA 190
 QY 183 MLEETSSNGGOSTVOQSTDGYFNNPFLGISTGNFTYLSNVND-----VYASL 231
 Db 191 MSLSLNNMNDNLNRGVQDSYVDTFPNO-----VTLTINIRNEPIVDSLHPYAVL 240
 QY 232 AITMLFVC 238
 Db 241 AITMLFVC 247

						RESULT 13
ID	Q38760	PRELIMINARY:	PRT:	252 AA.		
AC	Q38760:					
DT	01-NOV-1998 (TrEMBLrel. 08, Created)					
DR	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)					
DI	-1-PBC-2001 (TrEMBLrel. 19, Last annotation update)					
DE	Abrin-E (rRNA N-glycosidase) (EC 3.2.2.22) (Fragment).					
DN	Abrin-E (rRNA N-glycosidase) (EC 3.2.2.22) (Fragment).					
GN	RIP.					
OS	Adonis precatorius (Indian licorice) (Crab's eye).					
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
CC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;					
CC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Abreae; Adrus.					
OX	NCHI_TaxID=3816;					
RN	(1)					
RP	SEQUENCE FROM N.A.					
RC	TISSEUE=LEAF;					
RA	MEDLINE=91201329; Pubmed=2016300;					
RT	Evensen G., Mathiesen A., Sundan A.;					
FT	"Direct molecular cloning and expression of two distinct abrin A-					
RL	chains.";					
J.	Biol. Chem. 266:6848-6852(1991).					
CC	-1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN					
CC	SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL					
CC	SUBUNIT BY REMOVING ADENINE FROM POSITION 4,124 OF 28 S RNA.					
CC	ABRIN-A IS MORE TOXIC THAN RICIN.					
CC	-1- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT					
CC	FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT					
CC	PRECEDES ENOCYTOSIS.					
CC	-1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE					
CC	SPECIFIC ADENOSINE ON THE 28S RNA.					
CC	-1- SUBUNIT: DISULPHIDE-LINKED DIMER OF A AND B CHAINS.					
CC	-1- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN					
CC	CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).					
CC	-1- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING					
CC	PROTEINS. BELONGS TO TYPE 2 RIP.					
DR	EMBL: X54872; CAA38654.1..					
DR	HSSP: P11140; IABR.					
DR	ZincPro: IPB001574; RIP.					
DR	Pfam: PF00161; RIP. 1					
DR	PRINTS: PR00366; SHIGARICIN.					
KW	Hydroxylase; Glycosylase; Toxin; Repeat; Glycoprotein; Lectin.					
FT	CHAIN 1 252					
FT	NONTER 252 252					
SO	SOURCE 252 AA; BBFC94689592B5DE CRC64;					
Query Match 32.1%; Score 458; DB 10; Length 252;						
Best Local Similarity 45.6%; Pad. No. 8,5e-34;						
Matches 108; Conservative 34; Mismatches 91; Indels 4; Gaps 4;						
QY	4 TGDGFFETTLIRRVYSGSFNSNEIPLRNOSTIEVSDAQRFVLVELFMQGDDSTTAIAIDV 63					
DB	14 TDSQYKQPIELAIRRL-RGLGIHLIDIFVLDLP-TVEERNRYIVELISEEBSLETVIDV 71					
QY	64 TNYLVVAQAQDGQGYFLRDAPDAERHNLFTGTTRSSLPTFGSYTDLERYAGH-RDOIPLG 122					
DB	72 TNAYVVAAVASGSQYFLRDAPASASTYLEFTGTORSLRFDGSGYGLDERNAQTREOISLG 131					
QY	123 IELLGVSAARYGGSGTRAQASLLILLOMISAARFNFIFFRAAOYLNGSGEFLLPDAT 182					
DB	132 LGQLTHAISPFGSGSNDEKAKTTLLIVIOIAASEAARYRYSINRWGVSIRGTAFOPDPA 191					
QY	183 MELELSWGQGSTVOGOSTDGVANNPFRL-GTSGNGEVLTASNVRDVIASLIAMLFVC 238					
DB	192 MELSLNNMDNISRGVQSVDTPPNAYTLRRYNNQPIYVDSLHQSVAVLATMLFVC 248					
RESULT 14						
ID	Q6MZ67	PRELIMINARY:	PRT:	592 AA.		

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 15:20:00 : Search time 2037.82 Seconds
(without alignments)
11953.474 Million cell updates/sec

Title: US-09-627-165E-15

Perfect score: 837
Sequence: 1 catcacagcagcagcgagcagca.....aaccacccgtgcacatgta 837

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_pl: *
8: gb_pl: *
9: gb_pl: *
10: gb_pl: *
11: gb_pl: *
12: gb_pl: *
13: gb_pl: *
14: gb_pl: *
15: gb_pl: *
16: gb_pl: *
17: gb_pl: *
18: gb_pl: *
19: gb_pl: *
20: gb_pl: *
21: gb_pl: *
22: gb_pl: *
23: gb_pl: *
24: gb_pl: *
25: gb_pl: *
26: gb_pl: *
27: gb_pl: *
28: gb_pl: *
29: gb_pl: *
30: gb_pl: *
31: gb_pl: *
32: gb_pl: *
33: gb_pl: *
34: gb_pl: *
35: gb_pl: *
36: gb_pl: *
37: gb_pl: *
38: gb_pl: *
39: gb_pl: *
40: gb_pl: *
41: gb_pl: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	826.4	98.7	1873	8	AF369961 Viscum al
2	695.2	83.1	1923	6	AS8957 Sequence 1
3	695.2	83.1	1923	6	AR164305 Sequence
4	692	82.7	1596	6	AX019438 Sequence
5	688.8	82.3	1602	8	AY081149 Viscum al
6	683.8	81.7	1598	6	AX019435 Sequence
7	677.4	80.9	1598	6	AY081148 Viscum al
8	668.4	79.9	1598	6	AX139571 Sequence
9	668.4	79.9	1598	6	AF508914 Viscum al
10	643.2	77.1	1774	6	AR164303 Sequence
11	643.2	77.1	1774	6	AR164303 Sequence
12	641	76.6	1763	6	AX019436 Sequence
13	640.2	76.5	1763	6	AX019436 Sequence
14	632.8	75.6	1762	6	AX139573 Sequence
15	631.6	75.5	1762	6	AX019439 Sequence
16	628.2	75.1	1762	6	AR164307 Sequence
17	620.4	74.1	1747	8	AF508915 Viscum al
18	614	73.4	1768	6	AX019440 Sequence
19	588.6	70.3	1768	6	AX139575 Sequence
20	579.4	69.2	1768	6	AF508916 Viscum al
21	572	62.4	1596	6	AX139585 Sequence
22	274.4	32.8	1596	6	AX019441 Sequence
23	246	29.4	1762	6	AX019442 Sequence
24	246	29.4	1762	6	AX019443 Sequence
25	230	27.5	2010	8	AX139583 Sequence
26	137.4	16.4	2010	8	AX039803 Cinnamomu
27	135.8	16.2	1932	8	AF259548 Cinnamomu
28	135.8	16.2	2034	8	AY039801 Cinnamomu
29	134.2	16.0	2070	8	AY039802 Cinnamomu
30	116.2	13.9	1768	8	AF161672 A. precatoriu
31	114.6	13.7	1584	8	ABRIABRD A. precatoriu
32	114.6	13.7	2743	8	AFAB A. precatoriu
33	114	13.6	1763	8	AFAB A. precatoriu
34	114	13.6	1763	8	AFAB A. precatoriu
35	113.2	13.5	1584	8	ABRIABRD A. precatoriu
36	112.4	13.4	804	6	E09666 DNA encodin
37	112.4	13.4	753	8	APR167L A. precatoriu
38	112.4	13.4	753	8	APR167L A. precatoriu
39	111.6	13.3	804	6	E09665 DNA encodin
40	111.6	13.3	804	6	E09665 DNA encodin
41	110.8	13.2	753	8	APR167L A. precatoriu
42	98.6	11.8	1581	8	ABRIABRD A. precatoriu
43	93.4	11.2	1873	8	SN027122 Sambucus n1
44	93	11.1	1858	8	SN058357 Sambucus n1
45	89.4	10.7	1892	8	SRB400822 Sambucus n1

ALIGNMENTS

RESULT 1	AF369961	1873 bp	DNA	linear	PLN 12-DEC-2001
LOCUS	AF369961	1873 bp	DNA	linear	PLN 12-DEC-2001
DEFINITION	Viscum album subsp. coloratum VCA precursor, gene, complete cds.				
ACCESSION	AF369961				
VERSION	AF369961.1	GI:17529700			
KEYWORDS	Viscum album subsp. coloratum, Viscum album subsp. coloratum, Eukaryota, Viridiplantae, Streptophyta, Santalales, Viscaceae, Viscum.				
SOURCE	Viscum album subsp. coloratum, Viscum album subsp. coloratum, Eukaryota, Viridiplantae, Streptophyta, Santalales, Viscaceae, Viscum.				
ORGANISM	Viscum album subsp. coloratum, Viscum album subsp. coloratum, Eukaryota, Viridiplantae, Streptophyta, Santalales, Viscaceae, Viscum.				
REFERENCE	1 (bases 1 to 1873)				
AUTHORS	Park, W.-B. and Lyu, S.				
TITLE	Cloning of Viscum album subsp. coloratum (Korean mistletoe)				

JOURNAL Biochem. Biophys. Res. Commun. (2002) In press
REFERENCE 2 (bases 1 to 1873)
AUTHORS Park, W.-B. and Iyu, S.
TITLE Direct Submission
JOURNAL Submitted (13-APR-2001) Chemistry, Seoul Women's University, 126,
Kongnung-2dong, Nowon-gu, Seoul 139-774, Korea
FEATURES
source Location/Qualifiers
1..1873
/organism="Viscum album subsp. coloratum"
/sub_species="coloratum"
/db_xref="taxon:159976"
<138..>1835
/product="VCA precursor"
138..1835
/codon_start=1
/product="VCA precursor"
/protein_id="F61404.1"
/db_xref="GI:17529701"
/translation="MNLASRRAMVSCFLMLGLVGAATYKATKRYERLRATYTHO
TTGDEYFETPLRDYVSSGSFSENEIPILROSTIPVSDAFREVLVETWOGSDSTTA
IDVTLVYVAQADQSYFLRDAPEGRHLFTGTTRSLPEFTGSTDLEARYGRDQ
IPGLIEELIOSALRYSQSTRAQARSLILIQMISEAREPILWRARQYINSGES
FLPDMYLEETSWGQSTQVOQSTQDVENNPFRIGISTENFVLSNRDCAVDSLAIM
LVCMDRPSSDVAYWPLYIRVLENSGAVDVCTSEPTVYVIRGDLICADVDSLAK
FYNGNPDIOLMPDNPQIWTIRDTIGSGHCLITTYGTAGVYVLEFDCNVRATL
WQIMNGNTIIPRSNLVLAAGSSGTTLTQVQVYFLGQMLAGNDPAPEVYIYGF
GNLCHEANGASVETGCGSKENOMALYGNISIRPKNODCLTSGDSVSTENIV
SCSAGSGORWEFTNEGTILNLNGLVMDVAQSNPSLRILIIYPATGKNQMLPVP"

sig_peptide 138..203
mat_peptide 204..956
misc_feature 957..1061
misc_feature /note="Region: linker peptide"
mat_peptide /product="VCA beta chain"
BASE COUNT 456 a 458 c 492 g 467 t
ORIGIN

Query Match 98.7%; Score 826.4; DB 8; Length 1873;
Best Local Similarity 99.3%; Pred. No. 3,7e-237;
Matches 830; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CATGAGACGAGCGGCGACGATATATTCGGTTCATCAGCTTCTCCGAGATTATGTCGA 60
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 264 CATCAACACAGCGGCGACGATATATTCGGTTCATCAGCTTCTCCGAGATTATGTCGA 323
QY 61 AGCGGAGCTTTCCAAATGAGATACCACTGTGTCAGTACAGTCCCGTCGGAT 120
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 324 AGCGGAGCTTTCCAAATGAGATACCACTGTGTCAGTACAGTCCCGTCGGAT 383
QY 121 GCGCAAGATTGTGTGTGAGACTCACCATATGAGGGGAGACTCGATACAGCGCC 180
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 384 GCGCAAGATTGTGTGTGAGACTCACCATATGAGGGGAGACTCGATACAGCGCC 443
QY 181 ATCGAGCTTACCAATCTGTACGTGTGCTTACCAAGACGACCAACCTACTTTT 240
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 444 ATCGAGCTTACCAATCTGTACGTGTGCTTACCAAGACGACCAACCTACTTTT 503
QY 241 GCGACGACACGAGCGCGGGAAGGATCTTTCACGCGACACACAGATCCCTCTC 300
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 504 GCGACGACACGAGCGCGGGAAGGATCTTTCACGCGACACACAGATCCCTCTC 563
QY 301 CCATTTCAGGAGCTACACAGATCTGGAGGATACGCGGATATGGAGACAGATCCCT 360
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 564 CCATTTCAGGAGCTACACAGATCTGGAGGATACGCGGATATGGAGACAGATCCCT 623
QY 361 CTGGGTTATAGAGCACTCAATTCATTCGCTCTCGCGCTTCGTTATTCACAGGGGACGAC 420
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 624 CTGGGTTATAGAGCACTCAATTCATTCGCTCTCGCGCTTCGTTATTCACAGGGGACGAC 683
QY 421 CGGGCCCAAGCTCTCTTCTTAACTCTCAATTCAGATGATCTCCGAGGCGCGAGATTC 480
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 684 CGGGCCCAAGCTCTCTTCTTAACTCTCAATTCAGATGATCTCCGAGGCGCGAGATTC 743

QY 481 AATCCCATCTTTTGGAGGCTCGCCATACATTAACAGCGGAGATCTTTCTCCGAC 540
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 744 AATCCCATCTTTTGGAGGCTCGCCATACATTAACAGCGGAGATCTTTCTCCGAC 803
QY 541 ATGTCATGCTCGAGCTGAGACACTAGTTGGGCGCAACATCCAGCAGTCAGACT 600
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 804 ATGTCATGCTCGAGCTGAGACACTAGTTGGGCGCAACATCCAGCAGTCAGACT 863
QY 601 ACGGATGCGCTTTTAAATACCATTTTCGTTGGGATATATCCAGCGTACTTGACG 660
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 864 ACGGATGCGCTTTTAAATACCATTTTCGTTGGGATATATCCAGCGTACTTGACG 923
QY 661 TTGAGCATGTTGCGACAGCTGATCCGAGCTTACGATCATGTTGTTGATAGAGGAC 720
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 924 TTGAGCATGTTGCGACAGCTGATCCGAGCTTACGATCATGTTGTTGATAGAGGAC 983
QY 721 CGACCATCTTCGCGACAGCTGATCCGAGCTTACGATCATGTTGTTGATAGAGGAC 780
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 984 CGACATCTTCTCCAGCTGAGCTGATCCGAGCTTACGATCATGTTGTTGATAGAGGAC 1043
QY 781 ACGGCGCGCTCGACAGATTTACTGACAGCTTCCGACCCAGCGTCCATCGT 836
||||| (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||) (|||||)
Db 1044 ACGGCGCGCTCGACAGATTTACTGACAGCTTCCGACCCAGCGTCCATCGT 1099

RESULT 2
LOCUS A58957 1923 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 1 from Patent WO9701636.
ACCESSION A58957
VERSION A58957.1 GI:3714428
KEYWORDS
SOURCE
ORGANISM
Viscum album.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
Santalales; Viscaceae; Viscum.
1 (bases 1 to 1923)
Lentzen, H., Eck, U., Baur, A. and Zinke, H.
Recombinant mistletoe lectin (rML)
Patent: WO 9701636-A 1 16-JAN-1997;
MADANS AG KOEHLN (DE)
FEATURES
source Location/Qualifiers
1..1923
/organism="Viscum album"
/db_xref="taxon:3972"
/tissue_type="LEAF"
55..1749
/product="1"
/codon_start=1
/protein_id="CAA03513.1"
/db_xref="GI:3714428"
/translation="MNAVMDSRAMVSCFLMLGLVGAATYKATKRYERLRATYTHO
TTGDEYFETPLRDYVSSGSFSENEIPILROSTIPVSDAFREVLVETWOGSDSTTA
IDVTLVYVAQADQSYFLRDAPEGRHLFTGTTRSLPEFTGSTDLEARYGRDQ
IPGLIEELIOSALRYSQSTRAQARSLILIQMISEAREPILWRARQYINSGES
FLPDMYLEETSWGQSTQVOQSTQDVENNPFRIGISTENFVLSNRDCAVDSLAIM
LVCMDRPSSDVAYWPLYIRVLENSGAVDVCTSEPTVYVIRGDLICADVDSLAK
FYNGNPDIOLMPDNPQIWTIRDTIGSGHCLITTYGTAGVYVLEFDCNVRATL
WQIMNGNTIIPRSNLVLAAGSSGTTLTQVQVYFLGQMLAGNDPAPEVYIYGF
GNLCHEANGASVETGCGSKENOMALYGNISIRPKNODCLTSGDSVSTENIV
SCSAGSGORWEFTNEGTILNLNGLVMDVAQSNPSLRILIIYPATGKNQMLPVP"
55..153
sig_peptide
154..900
mat_peptide
901..957
mat_peptide /product="A-CHAIN"
938..1746
mat_peptide /product="PROPEPTIDE"
BASE COUNT 465 a 488 c 495 g 475 t
ORIGIN

Query Match 83.1%: Score 695.2; DB 6; Length 1923;
Best Local Similarity 90.7%: Pred. No. 1.1e-197;
Matches 758: Conservative 0; Mismatches 63; Indels 15; Gaps 1;

1 CATCAGACGAGCGGGAGCAATATTCCTGATCAGCGTTCCTCCAGATATATGTCTCA 60
|||||
181 CATCAAAACCGGGAGCAATATTCCTGATCAGCGTTCCTCCAGATATATGTCTCA 240
|||||
61 AGCGAAGCTTTTCCATGATGATACACTCTGCGTACGATCAGATCCCGTCTGAT 120
|||||
241 AGCGAAGCTTTTCCATGATGATACACTCTGCGTACGATCAGATCCCGTCTGAT 300
|||||
121 GCGCAAAAGATTTGTGTGTGATGATACCAATCAAGGGGGAGACTGATCAGCGCGCC 180
|||||
301 GCGCAAAAGATTTGTGTGTGATGATACCAATCAAGGGGGAGACTGATCAGCGCGCC 360
|||||
181 ATGAGGTACCAATCTGTAGTGTGCTTACCAAGACGCGCAATCTCTTTTG 240
|||||
361 ATGAGGTACCAATCTGTAGTGTGCTTACCAAGACGCGCAATCTCTTTTG 420
|||||
241 GCGAGCGACACAGACGCGCGGAAAGCATCTCTACCGCGCACCCAGATCTCTC 300
|||||
421 GCGAGCGACACAGACGCGCGGAAAGCATCTCTACCGCGCACCCAGATCTCTC 480
|||||
301 CCATTACCGAAGCTTACACAGATCTGAGACGATACGCGGTCATAGGACAGATCCCT 360
|||||
481 CCATTACCGAAGCTTACACAGATCTGAGACGATACGCGGTCATAGGACAGATCCCT 540
|||||
361 CTGGGTATAGAGAACTCATTCATCCGTCGCGGCTTCGTTATCCAGCGCGCACACC 420
|||||
541 CTGGGTATAGAGAACTCATTCATCCGTCGCGGCTTCGTTATCCAGCGCGCACACC 600
|||||
421 CGGGCCCAAGCTCTGTCCTTATATCCATTCAGATGATCTCCGAGCGCGAGATTC 480
|||||
601 CGTACCAAGCTCTGTCCTTATATCCATTCAGATGATCTCCGAGCGCGAGATTC 660
|||||
481 AATCCCATCTTTTGGAGGGCGGCAATATTAACAGCGGAGTATTTCTCCGAC 540
|||||
661 AATCCCATCTTTTGGAGGGCGGCAATATTAACAGCGGAGTATTTCTCCGAC 720
|||||
541 ATGTACATGCTGAGACTGAGACTAGTTGGGGCAACATCCAGCAATGCCAGAGTCT 600
|||||
721 GTGTACATGCTGAGACTGAGACTAGTTGGGGCAACATCCAGCAATGCCAGAGTCT 780
|||||
601 ACGGATGCGGTTTTTAAATTAACCATTTTCGTTGGGTATATCCACCGGTAACTCTGAGC 660
|||||
781 ACGGATGCGGTTTTTAAATTAACCATTTTCGTTGGGTATATCCACCGGTAACTCTGAGC 840
|||||
661 TTGAGCAATGTTCCGAGCGTATGAGCGATGAGATGTTGTTGATATAGGAGC 720
|||||
841 TTGAGCAATGTTCCGAGCGTATGAGCGATGAGATGTTGTTGATATAGGAGC 900
|||||
721 CGACCAATCTTCCGAGCGTATGAGCGATGAGCGATGAGATGTTGTTGATATAGGAGC 780
|||||
901 CGGCAATCTTCTCTGAGAGTCCGCTATTTGGCGCTATGATACAGCCCGATGAT 953
|||||
781 AGCGGCGCGGTGAGAGATGTTACTGATGCTTCGAGACCCAGCGGATGAT 836
|||||
954 -----AGCGATGATGTTACTGATGCTTCGAGACCTACGAGTGGGATTTG 1001
|||||

RESULT 3
AR164305 1923 bp DNA linear PAT 17-OCT-2001
LOCUS AR164305
DEFINITION Sequence 34 from patent US 6271368.
ACCESSION AR164305
VERSION AR164305.1 GI:16235409
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Lentzen, H., Eck, J., Baur, A. and Zinke, H.

TITLE Recombinant mistletoe lectin (rML)
JOURNAL Patent: US 6271368-A 34 07-AUG-2001;
FEATURES Location/Qualifiers
source 1.1923
BASE COUNT 465 a 488 c 495 g 475 t
ORIGIN

Query Match 83.1%: Score 695.2; DB 6; Length 1923;
Best Local Similarity 90.7%: Pred. No. 1.1e-197;
Matches 758: Conservative 0; Mismatches 63; Indels 15; Gaps 1;

1 CATCAGACGAGCGGGAGCAATATTCCTGATCAGCGTTCCTCCAGATATATGTCTCA 60
|||||
181 CATCAAAACCGGGAGCAATATTCCTGATCAGCGTTCCTCCAGATATATGTCTCA 240
|||||
61 AGCGAAGCTTTTCCATGATGATACACTCTGCGTACGATCAGATCCCGTCTGAT 120
|||||
241 AGCGAAGCTTTTCCATGATGATACACTCTGCGTACGATCAGATCCCGTCTGAT 300
|||||
121 GCGCAAAAGATTTGTGTGTGATGATACCAATCAAGGGGGAGACTGATCAGCGCGCC 180
|||||
301 GCGCAAAAGATTTGTGTGTGATGATACCAATCAAGGGGGAGACTGATCAGCGCGCC 360
|||||
181 ATGAGGTACCAATCTGTAGTGTGCTTACCAAGACGCGCAATCTCTTTTG 240
|||||
361 ATGAGGTACCAATCTGTAGTGTGCTTACCAAGACGCGCAATCTCTTTTG 420
|||||
241 GCGAGCGACACAGACGCGCGGAAAGCATCTCTACCGCGCACCCAGATCTCTC 300
|||||
421 GCGAGCGACACAGACGCGCGGAAAGCATCTCTACCGCGCACCCAGATCTCTC 480
|||||
301 CCATTACCGAAGCTTACACAGATCTGAGACGATACGCGGTCATAGGACAGATCCCT 360
|||||
481 CCATTACCGAAGCTTACACAGATCTGAGACGATACGCGGTCATAGGACAGATCCCT 540
|||||
361 CTGGGTATAGAGAACTCATTCATCCGTCGCGGCTTCGTTATCCAGCGCGCACACC 420
|||||
541 CTGGGTATAGAGAACTCATTCATCCGTCGCGGCTTCGTTATCCAGCGCGCACACC 600
|||||
421 CGGGCCCAAGCTCTGTCCTTATATCCATTCAGATGATCTCCGAGCGCGAGATTC 480
|||||
601 CGTACCAAGCTCTGTCCTTATATCCATTCAGATGATCTCCGAGCGCGAGATTC 660
|||||
481 AATCCCATCTTTTGGAGGGCGGCAATATTAACAGCGGAGTATTTCTCCGAC 540
|||||
661 AATCCCATCTTTTGGAGGGCGGCAATATTAACAGCGGAGTATTTCTCCGAC 720
|||||
541 ATGTACATGCTGAGACTGAGACTAGTTGGGGCAACATCCAGCAATGCCAGAGTCT 600
|||||
721 GTGTACATGCTGAGACTGAGACTAGTTGGGGCAACATCCAGCAATGCCAGATTC 780
|||||
601 ACGGATGCGGTTTTTAAATTAACCATTTTCGTTGGGTATATCCACCGGTAACTCTGAGC 660
|||||
781 ACGGATGCGGTTTTTAAATTAACCATTTTCGTTGGGTATATCCACCGGTAACTCTGAGC 840
|||||
661 TTGAGCAATGTTCCGAGCGTATGAGCGATGAGATGTTGTTGATATAGGAGC 720
|||||
841 TTGAGCAATGTTCCGAGCGTATGAGCGATGAGATGTTGTTGATATAGGAGC 900
|||||
721 CGACCAATCTTCCGAGCGTATGAGCGATGAGCGATGAGATGTTGTTGATATAGGAGC 780
|||||
901 CGGCAATCTTCTCTGAGAGTCCGCTATTTGGCGCTATGATACAGCCCGATGAT 953
|||||
781 AGCGGCGCGGTGAGAGATGTTACTGATGCTTCGAGACCCAGCGGATGAT 836
|||||
954 -----AGCGATGATGTTACTGATGCTTCGAGACCTACGAGTGGGATTTG 1001
|||||

RESULT 4
AX019438 1596 bp DNA linear PAT 07-SEP-2000
LOCUS AX019438
DEFINITION Sequence 15 from Patent W09940109.

```

ACCESSION  AX019438
VERSION     AX019438.1  GI:10043377
KEYWORDS
SOURCE      synthetic construct.
            artificial sequences.
            1 (bases 1 to 1596)
REFERENCE   Weilers, P., Stiefel, T., Voelker, W. and Morris, P.
TITLE       Recombinant mistletoe lectines
JOURNAL     Patent: WO 94/0109-A 15 12-AUG-1999;
            WELTERS, PETER (DE); STIEFEL, THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH
            (DE); VOELKER, WOLFGANG (DE); MORRIS, PETER (GB)
FEATURES
  source     1..1596
             /organism="synthetic construct"
             /db_xref="taxon:32630"
BASE COUNT  387 a 418 c 419 g 372 t
ORIGIN
Query Match      82.7%; Score 692; DB 6; Length 1596;
Best Local Similarity 90.4%; Pred. No. 9, 6e-197;
Matches 756; Conservative 0; Mismatches 65; Indels 15; Gaps 1;

OY 1  CATCAGACGAGCGGCGCAGATATTTCCGGTTCATCAGCCTTCTCCGAGATTATGTCTCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 28  CATCAACACGAGCGGCGAGGAATACCTCGGTCATCAGCCTTCTCCGAGATTATGTCTCA 87
OY 61  AGCGAAGCTTTTCCAAATGAGATACCACTCTGCGTCAGTCTACAGATCCCGTCCGAT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 88  AGCGAAGCTTTTCCAAATGAGATACCACTCTGCGTCAGTCTACAGATCCCGTCCGAT 147
OY 121  GCGCAAGATTGTGTGTGTGAGACTCACCACATCAGGCGGAGAGACTCGATCAGCGCGCC 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 148  GCGCAAGATTGTGTGTGTGAGACTCACCACATCAGGCGGAGAGACTCGATCAGCGCGCC 207
OY 181  ATGCACTTTCGAATCTGACGTGTGCTTACCAACAGCGAGACCATCTACTTTTG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 208  ATCAGCTTACCAATCTGACGTGTGCTTACCAACAGCGAGACCATCTACTTTTG 267
OY 241  CGCGAGCAGCAGCGGCGGAGAAAGCATCTCTTACCGGACACCAAGATCCTCTGC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 268  CGCGAGCAGCAGCGGCGGAGAAAGCATCTCTTACCGGACACCAAGATCCTCTGC 327
OY 301  CCATTCAACCGGAGCTACAGATCTGAGAGATAGCGCGGTATAGGAGACAGATCCCT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 328  CCATTCAACCGGAGCTACAGATCTGAGAGATAGCGCGGTATAGGAGACAGATCCCT 387
OY 361  CTGGGTATAGGAATCTCATTCATCGCTCGCGGCTTGTATCCAGGCGGAGAGACC 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 388  CTGGGTATAGGAATCTCATTCATCGCTCGCGGCTTGTATCCAGGCGGAGAGACC 447
OY 421  CGGGCCCAAGCTCTCCCTTAAATCTCATTCAGATGATCCGAGAGCGGCGAGATTC 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 448  CGTACCCAGACTCTGTCATTTAACTCTCATTCAGATGATCCGAGAGCGGCGAGATTC 507
OY 481  AATCCATCTTTTGGAGGCTCGCCATATTAACAGCGGAGATCATTTCTTCCGAC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 508  AATCCATCTTTATGAGGCTCGCCATATTAACAGCGGAGATCATTTCTTCCGAC 567
OY 541  ATGTAATGCTCGAGCTGAGACTAGTTGGGCCAACAATCCAGCAAGTCCACAGCTCT 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 568  GTGTACATGCTGGAGCTGAGAGAGATTGGGGCAACAATCCAGCAAGTCCACAGCTCT 627
OY 601  ACGATGAGCTTTTAAATTAACCATTTGGTTGGTATATCAACGAGTATCTGTAGC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 628  ACGATGAGCTTTTAAATTAACCATTTGGTTGGTATATCAACGAGTATCTGTAGC 687
OY 661  TTGAGCAATGTCGAGCTGATCGCAGCTTACCAATCAAGTGTGTATAGTAGGAGC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 688  TTGAGCAATGTCGAGCTGATCGCAGCTTACCAATCAAGTGTGTATAGTAGGAGC 747
OY 721  CGACCAATCTCTCGAGCTGCGTATGTCGCGCTGCTATACAGCCGCTCTGGAAT 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 748  CGGCATCTTCTCTACGTCGGCTATTGGCCCGCTGCTATACAGCCGCTGAT----- 800
OY 781  AGCGGCGCGTGCAGCATGTTACTGCACTGCTTCCGAACCCAGCGTCCGATCCT 836
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 801  -----AGCGATGATGTTACTGCACTGCTTCCGAACCTACGTCGGAATTGT 848

RESULT 5
AY081149
LOCUS      AY081149
DEFINITION Viscum album lectin chain A isoform 1 precursor, mRNA, partial cds.
ACCESSION AY081149
VERSION   AY081149.1  GI:19526600
KEYWORDS
SOURCE     Viscum album.
            Viscum album.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Santalales; Viscaceae; Viscum.
            1 (bases 1 to 1602)
REFERENCE   Santales, Viscaceae; Viscum.
AUTHORS    Paramasivam, M., Misra, V., Srinivasan, A. and Singh, T. P.
TITLE       Viscum album (Indian) mRNA for Mistletoe lectin chain A isoform 1
            and chain B
JOURNAL     Unpublished
            2 (bases 1 to 1602)
REFERENCE   Paramasivam, M., Misra, V., Srinivasan, A. and Singh, T. P.
AUTHORS    Direct Submission
TITLE       Submitted (21-FEB-2002) Biophysics, All India Institute of Medical
            Sciences, Ansari Nagar, New Delhi 110029, India
JOURNAL
FEATURES
  source     1..1602
             /organism="Viscum album"
             /db_xref="taxon:3972"
             /rname="Viscum album"
             /country="India"
             <1..1596
             /note="toxic component"
             /product="lectin chain A isoform 1 precursor"
             /codon_start=1
             /protein_id="AAB67006.1"
             /db_xref="GI:19526601"
             /translation="YERLRIRVHTGTEERYFRFTLLRDYVSSGSFSENEIPLRQST
             IFSDAOEFVVELJTEGDSITTAIDVTNLYVAYVAGDSYFLDARARAEHLPT
             GTTRESLPNGSYPDLEKRYAGHRDQIPGIDQLQSVTALRFPFGQRTQASLILIL
             QMISEANFENILMRARQYINGASFLPDVYMLLETSMGDSQTOVOSTEGBVFNPI
             RLAIPEGFTLVNVRVLAIALMLFVCGRSPSSDVRWPIVIRPVYIADVTCASL
             EPTVRIVGRNCVYRDDDFHDQIOLAMPSSNDPNOLWTIKRSDTINSQCLT
             TYGTAGYVMIIFDCNTAVREATIMQINGSTIIPNSIALAASGIGTTLIVOTL
             DYTGGWLAENDTAPREVTIYGFRLCMESNGSVYETCVSSQDNQMALYGDGSI
             RKRONQDCLTCGRDSYVYINIVSCSAGSGORWVPTNIGAILINKGLAMDVGSN
             PKRLIITIPATGKPNQMLPVP"
             <1..747
             /product="lectin chain A isoform 1"
             /note="toxic component; ribosome-inactivating protein"
             /note="toxic component; ribosome-inactivating protein"
             748..804
             misc_feature
             /note="Region: linker peptide"
             mat_peptide
             /product="lectin B chain"
             805..1593
             3'UTR
             1597..1602
BASE COUNT  394 a 415 c 419 g 374 t
ORIGIN
Query Match      82.3%; Score 688.8; DB 8; Length 1602;
Best Local Similarity 90.2%; Pred. No. 8, 8e-196;
Matches 754; Conservative 0; Mismatches 67; Indels 15; Gaps 1;

OY 1  CATCAGACGAGCGGCGCAGATATTTCCGGTTCATCAGCCTTCTCCGAGATTATGTCTCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 28  CATCAACACGAGCGGCGAGGAATACCTCGGTCATCAGCCTTCTCCGAGATTATGTCTCA 87
OY 61  AGCGAAGCTTTTCCAAATGAGATACCACTCTGCGTCAGTCTACAGATCCCGTCCGAT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 88  AGCGAAGCTTTTCCAAATGAGATACCACTCTGCGTCAGTCTACAGATCCCGTCCGAT 147

```


Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 REFERENCE 1 (bases 1 to 747)
 AUTHORS Paramasivam, M., Srihivasan, A. and Singh, T.P.
 TITLE Viscum Album (Indian) mRNA for Mistletoe Lectin Chain A, Isoform 2
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 747)
 AUTHORS Paramasivam, M., Srihivasan, A. and Singh, T.P.
 TITLE Direct Submission
 JOURNAL Submitted (21-FEB-2002) Biophysics, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India

FEATURES
 source
 1..747
 /organism="Viscum album"
 /db_xref="taxon:3972"
 /tissue="leaf"
 /country="India"
 <1..>747
 /codon_start=1
 /product="lectin chain A isoform 2 precursor"
 /protein_id="AA187005.1"
 /db_xref="GI:19526599"
 /translation="YERLRURYHOTGDEYFRFTLLRDYVSGSPSEIPLLRKOST
 IPVSDAQRFVVELNFGSDSYTAIDVNIYVAIYAGDSYFLDAPDGERLFT
 GTRSLPFTSYSDRYRAGHRDQIPGIEELIOSVALRYPGSGTRADASIIILI
 OMISEAARFNPFIWRVRDINSSESLPDMYMLETSWGSQSTQVQSTDSVFNNPF
 RLAISTGNFVTLISNVRDVIASLAIMLFVCGE"
 <1..747
 /product="lectin chain A isoform 2"
 /note="toxic component; ribosome-inactivating protein"
 ORIGIN
 BASE COUNT 176 a 204 c 188 g 179 t

Query Match 80.9%; Score 677.4; DB 8; Length 747;
 Best Local Similarity 96.4%; Pred. No. 2, 2e-192;
 Matches 693; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 CATGAGACGAGCGGCGACGAATATTCGGGTTTCATCAGCTTTCGGAGATTATGTCTCA 60
 DB 28 CATCAACACGAGGCGGCGACGAATATTCGGGTTTCATCAGCTTTCGGAGATTATGTCTCA 87
 OY 61 AGCGAAGCTTTTCCAAATGAGATACACTCTTGGTCACTACGATCCCGCTCTCGAT 120
 DB 88 AGCGAAGCTTTTCCAAATGAGATACACTCTTGGTCACTACGATCCCGCTCTCGAT 147
 OY 121 GCGCAAGATTTGTGGTGGAACTCAACAATCAGGGGGAGACTCGATCAGGGCGCC 180
 DB 148 GCGCAAGATTTGTGGTGGAACTCAACAATGAGGGGAGACTCGATCAGGGCGCC 207
 OY 181 ATGACGCTTACCAATCTGATGCTGGCTTACCAACAGCGACCAATCCACTTTTGG 240
 DB 208 ATGACGCTTACCAATCTGATGCTGGCTTACCAACAGCGACCAATCCACTTTTGG 267
 OY 241 GCGAGCAGACAGAGCGCGCGAAGGATCTCTTCCGCGACCAACAGATCTCTCTC 300
 DB 268 GCGAGCAGACAGAGCGCGCGAAGGATCTCTTCCGCGACCAACAGATCTCTCTC 327
 OY 301 CCATTTCACGGAAGCTACACAGATCTGAGGAGTACCGCGGTCAATAGGAGCCAGATCCCT 360
 DB 328 CCATTTCACGGAAGCTACACAGATCTGAGGAGTACCGCGGTCAATAGGAGCCAGATCCCT 387
 OY 361 CTGGGTTATAGAGACCTCAATCAATCCGTCGCGGCTTGGTATACAGGGCGGAGATTC 420
 DB 388 CTGGGTTATAGAGACCTCAATCAATCCGTCGCGGCTTGGTATACAGGGCGGAGATTC 447
 OY 421 CGGGCCCAAGCTCGTTCCCTTATATCTCTCATAGATGATCTCCGAGGCGCGAGATTC 480
 DB 448 CGGGCCCAAGCTCGTTCCCTTATATCTCTCATAGATGATCTCCGAGGCGCGAGATTC 507
 OY 481 AATCCCAATCTTTTGAGAGGCTCGCCAAATCAATTAACAGCGGGAGTCAATTTTCCGAC 540
 DB 508 AATCCCAATCTTTTGAGAGGCTCGCCAAATCAATTAACAGTGGGAGTCAATTTTCCGAGC 567

OY 541 ATGTACATGCTCGAGCGAGAGACTAGTTGGGGCCCAACATCCACGAAGTCCACAGTCT 600
 DB 568 ATGTACATGCTCGAGCGAGAGAGTGGGGCCCAACATCCACGAAGTCCACAGTCT 627
 OY 601 ACGGATGGCGTTTATATACCATTTCGGTGGTATATACACGCGTAACTGGTGACG 660
 DB 628 ACGGATGGCGTTTATATACCATTTCGGTGGTATATACACGCGTAACTGGTGACG 687
 OY 661 TTGAGCAATGTTGCGACAGTGAATCCGACTTACCAATCATGTTGTTGTATAGGGA 719
 DB 688 TTGAGCAATGTTGCGACAGTGAATCCGACTTACCAATCATGTTGTTGTATAGGGA 746

RESULT 8
 AX139571 762 bp DNA linear PAT 30-MAY-2001
 LOCUS AX139571
 DEFINITION Sequence 1 from Patent EPI074560.
 ACCESSION AX139571
 VERSION AX139571.1 GI:14275207
 KEYWORDS
 SOURCE Viscum album subsp. coloratum.
 ORGANISM Viscum album subsp. coloratum.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 REFERENCE 1 (bases 1 to 762)
 AUTHORS Kim, J., Song, S., Suh, B., Lee, K., Doo, M., Kwak, J., Song, B., Yoon, T.,
 Kang, T. and Park, C.
 TITLE Crude extract from Viscum album coloratum, and proteins and lectins
 JOURNAL isolated therefrom
 Patent: EP 1074560-A 1 07-FEB-2001;
 Mistle Biotech Co., Ltd. (KR)
 location/Qualifiers
 FEATURES
 source
 1..762
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 ORIGIN
 BASE COUNT 172 a 220 c 186 g 184 t

Query Match 79.9%; Score 668.4; DB 6; Length 762;
 Best Local Similarity 94.4%; Pred. No. 1, 1e-189;
 Matches 693; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 1 CATGAGACGAGCGGCGACGAATATTCGGGTTTCATCAGCTTTCGGAGATTATGTCTCA 60
 DB 28 CATCAACACGAGGCGGCGACGAATATTCGGGTTTCATCAGCTTTCGGAGATTATGTCTCA 87
 OY 61 AGCGAAGCTTTTCCAAATGAGATACACTCTTGGTCACTACGATCCCGCTCTCGAT 120
 DB 88 AGCGAAGCTTTTCCAAATGAGATACACTCTTGGTCACTACGATCCCGCTCTCGAT 147
 OY 121 GCGCAAGATTTGTGGTGGAACTCAACAATCAGGGGGAGACTCGATCAGGGCGCC 180
 DB 148 GCGCAAGATTTGTGGTGGAACTCAACAATCAGGGGGAGACTCGATCAGGGCGCC 207
 OY 181 ATGACGCTTACCAATCTGATGCTGGCTTACCAACAGCGACCAATCCACTTTTGG 240
 DB 208 ATGACGCTTACCAATCTGATGCTGGCTTACCAACAGCGACCAATCCACTTTTGG 267
 OY 241 GCGAGCAGACAGAGCGCGCGAAGGATCTCTTCCGCGACCAACAGATCTCTCTC 300
 DB 268 GCGAGCAGACAGAGCGCGCGAAGGATCTCTTCCGCGACCAACAGATCTCTCTC 327
 OY 301 CCATTTCACGGAAGCTACACAGATCTGAGGAGTACCGCGGTCAATAGGAGCCAGATCCCT 360
 DB 328 CCATTTCACGGAAGCTACACAGATCTGAGGAGTACCGCGGTCAATAGGAGCCAGATCCCT 387
 OY 361 CTGGGTTATAGAGACCTCAATCAATCCGTCGCGGCTTGGTATACAGGGCGGAGATTC 420
 DB 388 CTGGGTTATAGAGACCTCAATCAATCCGTCGCGGCTTGGTATACAGGGCGGAGATTC 447
 OY 421 CGGGCCCAAGCTCGTTCCCTTATATCTCTCATAGATGATCTCCGAGGCGCGAGATTC 480
 DB 448 CGGGCCCAAGCTCGTTCCCTTATATCTCTCATAGATGATCTCCGAGGCGCGAGATTC 507
 OY 481 AATCCCAATCTTTTGAGAGGCTCGCCAAATCAATTAACAGCGGGAGTCAATTTTCCGAC 540
 DB 508 AATCCCAATCTTTTGAGAGGCTCGCCAAATCAATTAACAGTGGGAGTCAATTTTCCGAGC 567

Db 448 CGTGGCCAAAGCTGCTTCTTATCATCTCATTCAGATGATCTCCGAGGCCCGCCAGATTC 507
QY 481 AATCCCATCTTTTGGAGGCGTCCGCAATACATTACAGCGGGAGATCTTCTCCGAC 540
Db 508 AATCCCATCTTATGAGGCGTCCGCAATACATTACAGCGGGAGATCTTCTCCGAC 567
QY 541 ATGATACATGCTGAGCTGAGACTAGTTGGGGCCCAACATCCAGCAATCCAGAGTCT 600
Db 568 AGCTACATCTCCAGCTGAGAGAGTGGGGCCCAACATCCAGCAATCCAGAGTCT 627
QY 601 ACGATGCGCTTTTATTAACCATTCGTTGGGTATATCCAGCGGTACTCTGAGAG 660
Db 628 ACGATGCGCTTTTATTAACCATTCGTTGGGTATATCCAGCGGTACTCTGAGAG 687
QY 661 TTGACCATCTTCCGACGATGATCCGACCTTACGATCATCTTTGATATGAGGAC 720
Db 688 TTGACCATCTTCCGACGATGATCCGACCTTACGATCATCTTTGATATGAGGAC 747
QY 721 CGACCATCTCTCTC 734
Db 748 CGGCATCTCTCTC 761

RESULT 9
AF508914 762 bp mRNA linear PLN 03-JUN-2002
LOCUS Viscum album subsp. coloratum lectin chain A isoform 1 mRNA,
DEFINITION partial cds.
ACCESSION AF508914
VERSION AF508914.1 GI:21314411
KEYWORDS
SOURCE Viscum album subsp. coloratum.
ORGANISM Viscum album subsp. coloratum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Santalales; Viscaceae; Viscum.
1 (bases 1 to 762)
Park, C.H., Lee, D.W., Kang, T.B., Lee, K.H., Yoon, T.J., Kim, J.B.,
Do, M.S., and Song, S.K.
CDNA cloning and sequence analysis of the lectin genes of the
Korean mistletoe (Viscum album coloratum)
JOURNAL Mol. Cells 12 (2), 215-220 (2001)
MEDLINE 21566752
PUBMED 11710524
REFERENCE 2 (bases 1 to 762)
AUTHORS Park, C.-H., Lee, D.-W., Kang, T.-B., Lee, K.-H., Yoon, T.-J.,
Kim, J.-B., Do, M.-S., and Song, S.-K.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2002) School of Bioscience and Food Technology,
Hanyang Univ., Hanyangae-eup Namdong-ri, Pohang, Kyungbook 791-708,
Korea

FEATURES
source Location/Qualifiers
1..762
/organism="Viscum album subsp. coloratum"
/sub_species="coloratum"
/db_xref="taxon:159976"
<1..>762
/note="lectin A1 chain; KML-A1"
/codon_start=1
/product="lectin chain A isoform 1"
/protein_id="AA046932.1"
/db_xref="GI:21314412"
/translation="YERLRATVTHQTHGDEYFPIILADYVSSGFSFNEPIILROST
IPVSDARFVLELITNOCGDSITAIIVTMYVAIQADQSTFLRDAFDEAERHFT
GTRRSILPTFGSYTDLERFAGRQDIFLGRBELIVSLRPGSNTFAQKRSHTILI
OMISEARFPIILMKROISSGSGFLPDYIILQLETSSGQSTOVYHSTQVFNPI
RLTISTGVFTLSNWDVIALSLMLCEDEDRSSS"
misc_feature
748..>762
/note="Region: linker"
BASE COUNT 172 a 220 c 186 g 184 t
ORIGIN
Query Match 79.9%; Score 668.4; DB 8; Length 762;

Best Local Similarity 94.4%; Pred. No. 1.1e-189;
Matches 693; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 CATTGACGAGCGGCGCAGATATTTCCGTTTCATCAGCTTCTCCGAGTATATGCTCA 60
Db 28 CATTGACGAGCGGCGCAGATATTTCCGTTTCATCAGCTTCTCCGAGTATATGCTCA 87
QY 61 AGCGAAGCTTTTCCATGATGATACCACTCTGCGTACGTACAGTCCGCTCGGAT 120
Db 88 AGCGAAGCTTTTCCATGATGATACCACTCTGCGTACGTACAGTCCGCTCGGAT 147
QY 121 GCGCAAAATTTGTTGTTGATGACATCCATATCAGGGGAGAGATCCATACGCGCC 180
Db 148 GCGCAAAATTTGTTGTTGATGACATCCATATCAGGGGAGAGATCCATACGCGCC 207
QY 181 ATGACGTTACCAATCTGACGTGCTGCTTACCAAGCAGGAGCAATCTACTTTTGG 240
Db 208 ATGACGTTACCAATCTGACGTGCTGCTTACCAAGCAGGAGCAATCTACTTTTGG 267
QY 241 GCGAGCAGCAGACGCGCGGAAAGCATCTCTTACCGGCGACAGCATCTCTCTC 300
Db 268 GCGAGCAGCAGACGCGCGGAAAGCATCTCTTACCGGCGACAGCATCTCTCTC 327
QY 301 CCATTACCGGAGCTACACAGATCTGAGGAGATACCGGCTATAGGAGACAGATCCCT 360
Db 328 CCATTACCGGAGCTACACAGATCTGAGGAGATACCGGCTATAGGAGACAGATCCCT 387
QY 361 CTGGCTTATAGGAACTCATTCATCCGTCGCGGCTGCTTATCCAGGGGAGGAGAC 420
Db 388 CTGGCTTATAGGAACTCATTCATCCGTCGCGGCTGCTTATCCAGGGGAGGAGAC 447
QY 421 CGGCGCCCAAGCTGTTCCCTTATTAATCTCATTCAGTATGATCTCCGAGCGCGAGATTC 480
Db 448 CGTCCCGCAAGCTGTTCCCTTATTAATCTCATTCAGTATGATCTCCGAGCGCGAGATTC 507
QY 481 AATCCCATCTTATGAGGCGTCCGCAATACATTACAGCGGGAGATCTTCTCCGAC 540
Db 508 AATCCCATCTTATGAGGCGTCCGCAATACATTACAGCGGGAGATCTTCTCCGAC 567
QY 541 ATGATACATGCTGAGCTGAGACTAGTTGGGGCCCAACATCCAGCAATCCAGAGTCT 600
Db 568 ACGTACATCTCCAGCTGAGAGAGTGGGGCCCAACATCCAGCAATCCAGAGTCT 627
QY 601 ACGATGCGCTTTTATTAACCATTCGTTGGGTATATCCAGCGGTACTCTGAGAG 660
Db 628 ACGATGCGCTTTTATTAACCATTCGTTGGGTATATCCAGCGGTACTCTGAGAG 687
QY 661 TTGACCATCTTCCGACGATGATCCGACCTTACGATCATCTTTGATATGAGGAC 720
Db 688 TTGACCATCTTCCGACGATGATCCGACCTTACGATCATCTTTGATATGAGGAC 747
QY 721 CGACCATCTCTCTC 734
Db 748 CGGCATCTCTCTC 761

RESULT 10
AR164303 774 bp DNA linear PAT 17-OCT-2001
LOCUS Sequence 30 from patent. US 6271368.
DEFINITION AR164303
ACCESSION AR164303.1 GI:16235405
VERSION AR164303.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 774)
AUTHORS Lentzen, H., Eck, J., Baur, A., and Zinke, H.
TITLE Recombinant mistletoe lectin (rML)
JOURNAL Patent: US 6271368-A 30 07-AUG-2001;
FEATURES Location/Qualifiers
1..774
/organism="unknown"

BASE COUNT	173	a	232	c	185	g	184	t
ORIGIN								

AUTHORS	Eck, J. and Schmidt, A.
TITLE	RECOMBINANT FISSION PROTEINS BASED ON RIBOSOME-INACTIVATING PROTEINS

Query Match	77.18;	Score 645.2;	DB 6;	Length 774;
-------------	--------	--------------	-------	-------------

Patent: WO 9829540-A 7 09-JUL-1998;

Matches 677; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

FEATURES

Location/qualifiers

QY 1 CATCAGACGAGCGAATAATTCGGTTCATCAGCTTCTCCGAGATTATGCTCA 60

```

/organism="Viscum album"

```

Db 34 CACCAGACCACGGTGAAGATAATTCCGGTTCATCACGCTCTCCGAGATTATGCTCA 93

```
/tissue_type="LEAF"
```

Oy	601	ACGAGTGGCGTTTATTAATACCATTGTGGTGCTATTCACACGGTACTTGTGAG	660
Dd	628	ACGAGTGGCGTTTATTAATACCATTGTGGTGCTATTCACACGGTACTTGTGAG	687
Oy	661	TTAGCACAATGTTGCACGCTGATGCCAGCTTACCACATCATGTTGTTATATAGAGAC	720
Dd	688	TTAGCAATGTTCCGACGTGATGCCAGCTTGGCATCATGTTGTTATATGGGAGAG	747
Oy	721	CGACCATCT 729	
Dd	748	CGGCGATCT 756	
 RESULT 12 AX019436 763 bp DNA linear PAT 07-SEP-2000 LOCUS Sequence 13 from Patent W09940109. DEFINITION AX019436 ACCESSION AX019436 VERSION AX019436.1 GI:10043375 KEYWORDS SOURCE . ORGANISM synthetic construct. REFERENCE synthetic construct. AUTHORS artificial sequences. TITLE 1 (bases 1 to 763). JOURNAL Walters P., Stiefel T., Voelter W. and Morris P. REMARKS Recombinant mistletoe lectines Patient NO 9940109-A.13.12-Aug-1999; Welten PEIER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB) Location/Qualifiers FEATURES source 1..763 /organism="synthetic construct" /db_xref="taxon:37630"			
misc_feature 163 a 215 c 175 g 170 t 40 others			
BASE COUNT 319			
ORIGIN			
Query Match 76.6%; Score 641; DB 6; Length 763;			
Best Local Similarity 89.1%; Pred. No. 1.9e-181;			
Matches 655; Conservative 37; Mismatches 42; Indels 1; Gaps 1;			
Oy	1	CATCACAGCAGCGGGGAGCAATTTTCGGGTTTCATCAACGGTCTCCGAGATTATGTCTCA	60
Dd	28	CATCAACACGAGGGGAGKAAATCTTCGGGTTTCATCAACGGTCTCCGAGATTATGTCTCA	87
Oy	61	AGCGGAAGCTTTTCCATAGATGAGAATACACGCTTGCTGACGTACATGCCGCTCGGAT	120
Dd	88	AGCGGAAGCTTTTCCATAGATGAGAATACACGCTTGCTGACGTACATGCCGCTCGGAT	147
Oy	121	GCCCAAAAGTTTGTGTGTGTGTGGAATCTACCAAATACAGGGGGAGACATCGATCAGGCGCC	180
Dd	148	GCCCAAAAGTTTGTGTGTGTGTGGAATCTACCAAACAGGGGGSRBACTCGATTACGGCGCC	207
Oy	181	ATCGACGTTAACCATCTGTACGATGATGAGCTTACCAAGACAGCGAACCATTCTACTTTTG	240
Dd	208	ATCGACGTTAACCATCTGTACGATGATGAGCTTACCAAGACAGCGAACCATTCTACTTTTG	267
Oy	241	CGGAGACGACGAGCGGGGAAAAGGACCTCTCTACCGGACACACAGA-TCTCTCTCT	299
Dd	268	CGGAGACGACGAGCGGGGAAAAGGACGACCTCTCTACCGGACACACAGATCTCTCTCT	327
Oy	300	CCCATTCACCGGAAGCTACACATATGTGAGGAGTAAGCCGGATATAGGAGACAGATCCC	359
Dd	328	CCCATTCACCGGAAGCTACACATATGTGAGGAGTAAGCCGGATATAGGAGACAGATCCC	387
Oy	360	TCGTGGTATAAGAGGACATCATCAATCCGCTTCGGCGCTTCGTTATTCACAGGGGACGAC	419
Dd	388	TCGTGGTATAAGAGGACATCATCAATCCGCTTCGGCGCTTCGTTATTCACAGGGGACGAC	447
Oy	420	CGGGGCCCAAGCTCGTTCCTTATTAATCTCATTCATGATGATCTCCGAGGCCGCGAGATT	479
Dd	448	GCGTGTCAAGCTCGTTCCTTATTAATCTCATTCATGATGATCTCCGAGGCCGCGAGATT	507

QY	480	CAATCCCATCTTTTGGAGGCTGGCCAAATACATTTAAACAGGCGGAGATTTCTTCCCGA	539
		: : : : : : : : :	
Db	508	CAATCCCATCTTTATGAGAGKATGCCCAKAVATTTAAAGTGGGGRCACTATTTCTGCGAGA	567
QY	540	CATGTAATGCTGCGAGCTGAGACATTAATTTGGGGGCAACAATCAGCAAGTCAGCAAGTC	599
		: : : : : : : : :	
Db	568	CGGTACATGCTGAGAGCTGAGACAGAGTTGGGGCCCAATCAGCAAGTCAGCAATTC	627
QY	600	TACGATGGCGCTTTTATATACCATTTTCGGTGGGATATTCACGGCTAACTTCGTCAG	659
		: : : : : : : : :	
Db	628	AACCGATGGCGCTTTTATATACCATTTTCGGTGGGATATTCACGGCTAACTTCGTCAG	687
QY	660	GTTGACCAATGTTGCGACAGCTGATGCCACAGCTTAGCGATCATGTTGTTGATAGGGA	719
		: : : : : : : : :	
Db	688	GTTGACCAATGTTGCGACAGCTGATGCCACAGCTTAGCGATCATGTTGTTGATAGGGA	747
QY	720	CGGACCATCTTCCTC	734
		: : : : : : : : :	
Db	748	CGGCGCATCTTCCTC	762
RESULT 13			
LOCUS	A90894	762 bp	DNA linear
DEFINITION	Sequence 1 from Patent WO9829540.		
ACCESSION	A90894		
VERSION	A90894.1	GI:6739419	
KEYWORDS			
SOURCE	European mistletoe.		
ORGANISM	Viscum album		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Sapnatales; Viscaceae; Viscum.		
REFERENCE	1 (bases 1 to 762).		
AUTHORS	Eck J, and Schmidt A.		
TITLE	RECOMBINANT FUSION PROTEINS BASED ON RIBOSOME-INACTIVATING PROTEINS		
JOURNAL	OP EUROPANANT MISTELTJEE VISCUM ALBUM		
	Patent: NO 9829540-A, 1.08.-JUN.-1998.		
FEATURES	ECK JDERGEN (DE); SCHMIDT ARNO (DE)		
source	Location/Qualifiers		
	1..762		
	/organism="Viscum album"		
	/db_xref="taxon:3572"		
	/tissue_type="LEAF"		
	4..>759		
	/function="TOXIC COMPONENT"		
	/note="FIG. 11A: NUKLEOTIDSEQUENZ UND ABGELEITETE AMINO		
	ACIDSEQUENZ"		
	/codon_start=1		
	/product="RECOMBINANT MISTELTJEE LECTIN A-CHAIN"		
	/protein_id="CAB69393.1"		
	/db_xref="GI:6739420"		
	/translation="MYERISILTAHQTEGEYFRFILLRDVYSSGFSNEIPLROS		
	TIPASDORFVLELVETNGGSGSLTAIDVTNLYVAYACGDSYFLDAPRAETHLE		
	TGTRRSFENSPYDLERYAGHRDIPGLIDGLLOSLVATLRPGSGTQARSLILE		
	IQMSEARFNPIILMAROVYINGASGLPDVYMLLETLSMOGOSTOVHSTDGVNNPD		
	IRLAIIPGNVLTIVRDVIAIIMLFCGRP"		
BASE COUNT	169 a	230 c	183 g 180 t
ORIGIN			
Query Match	76.5%:	Score 640.2:	DB 6: Length 762:
Best Local Similarity	92.7%:	Pred. No. 3.3e-181:	
Matches	672:	Conservative	0: Mismatches 53: Indels 0: Gaps 0:
QY	1	CATCAGACACGCGGAGACGATAATTTCCGGGTGATCAAGCTTCTCCGATATATGATCTCA	60
		: : : : : : : : :	
Db	34	CACACAGCACCGGAGAGATATTTCCGGTATCAAGCTTCTCCGATATATGATCTCA	93
QY	61	AGCGGAAGCTTTCCATGATGATACCACTCTGGGTCAAGTCACGATCCCGCTCGGAT	120
		: : : : : : : : :	
Db	94	AGCGGAAGCTTTCCATGATGATACCACTCTGGGTCAAGTCACGATCCCGCTCGGAT	153
QY	121	CGCGAAGATTTGTGTGGTGGAGACTCACCAATCAGGGGGGAGAGATCTGATACAGCGCGC	180

Db 154 GCGCAAGATTTGCTGTGTGAGCTCACCACAGGGGGAGACTCGATCCAGCGCC 213
 |||
 Oy 181 ATGACGTTACCAATCTGTAGTGTGGCTTACCAAGAGGGGACCAATCCACTTTTG 240
 |||
 Db 214 ATGACGTTACCAATCTGTAGTGTGGCTTACCAAGAGGGGACCAATCCACTTTTG 273
 |||
 Oy 241 GCGAGCGACAGAGCGCGGAAAGCATCTCTTACCGGACACCAAGATCCTCTCTC 300
 |||
 Db 274 GCGAGCGACAGAGCGCGGAAAGCATCTCTTACCGGACACCAAGATCCTCTCTC 333
 |||
 Oy 301 CCATTCACCGAAGCTACAGATCTGAGCGATACCGGTCATAGGAGCAGATCCCT 360
 |||
 Db 334 CCATTCACCGAAGCTACAGATCTGAGCGATACCGGTCATAGGAGCAGATCCCT 393
 |||
 Oy 361 CTGGATATAGAGAACTCATTCATCCGTCCGCGCTTGCTTATCCAGCGGAGCAGC 420
 |||
 Db 394 CTGGATATAGAGAACTCATTCATCCGTCCGCGCTTGCTTATCCAGCGGAGCAGC 453
 |||
 Oy 421 GCGGCCACAGCTGTTCCCTTAATCTCTCATTCAGATGATCTCGAGGCGCGAGATTC 480
 |||
 Db 454 CGTACCCAGCTGTTCCCTTAATCTCTCATTCAGATGATCTCGAGGCGCGAGATTC 513
 |||
 Oy 481 AATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGCGGAGATCATTTCTCCGAC 540
 |||
 Db 514 AATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGTGCGGCTCATTTCTGCCAGAC 573
 |||
 Oy 541 ATGTACATGCTGAGCTGAGACTAGTTGGGCGCCAAATCCAGCAATGTCAGAGTCT 600
 |||
 Db 574 GTGTACATGCTGAGCTGAGACTAGTTGGGCGCCAAATCCAGCAATGTCAGAGTCT 633
 |||
 Oy 601 ACGATAGCGGTTTTTAATACCATTTGCGTGGTATACCAAGCGTAACTCTGTCAG 660
 |||
 Db 634 ACGATAGCGGTTTTTAATACCATTTGCGTGGTATACCAAGCGTAACTCTGTCAG 693
 |||
 Oy 661 TTGAGCAATGTTGCGAGCTGATGCCAGCTTACGATGATGTTGTTATGAGGAC 720
 |||
 Db 694 TTGAGCAATGTTGCGAGCTGATGCCAGCTTACGATGATGTTGTTATGAGGAC 753
 |||
 Oy 721 CGACC 725
 |||
 Db 754 CGCCC 758
 |||

RESULT 14
 LOCUS AX139573 762 bp DNA linear PAT 30-MAY-2001
 DEFINITION Sequence 3 from Patent EP1074560.
 ACCESSION AX139573
 VERSION AX139573.1 GI:14275208
 KEYWORDS
 SOURCE
 ORGANISM
 Viscum album subsp. coloratum.
 Viscum album subsp. coloratum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 1 (bases 1 to 762)
 Kim,J., Song,S., Sub,B., Lee,K., Doo,M., Kwak,J., Song,B., Yoon,T.,
 Kang,T. and Park,C.
 Crude extract from Viscum album coloratum, and proteins and lectins
 isolated therefrom
 Patent: EP 1074560-A 3 07-FEB-2001;
 Mistle Biotech Co., Ltd. (KR)
 FEATURES
 location/Qualifiers
 1..762
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 BASE COUNT 173 a 225 c 179 g 184 t 1 others
 ORIGIN
 Query Match 75.6%; Score 632.8; Db 6; Length 762;
 Best Local Similarity 91.3%; Pred. No. 5.6e-179;
 Matches 670; Conservative 1; Mismatches 63; Indels 0; Gaps 0;

Oy 1 CATCAGACAGCGGCGAGATATTTCCGTTTCATCAGCTTCTCCAGATATATGTCGA 60
 |||
 Db 28 CATCAACACAGCGGCGAGCAATATTTCAATTCATCAGCTTCTCCAGATATATGTCGA 87
 |||
 Oy 61 AGCGAAGCTTTTCCATAGATACACCTTCCGTCAGTGTACAGATCCCGTCTGGAT 120
 |||
 Db 88 AGCGAAGCTTTTCCATAGATACACCTTCCGTCAGTGTACAGATCCCGTCTGGAT 147
 |||
 Oy 121 GCGCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180
 |||
 Db 148 AGCGAAGCTTTTCCATAGATACACCTTCCGTCAGTGTACAGATCCCGTCTGGAT 207
 |||
 Oy 181 ATGACGTTACCAATCTGTAGTGTGGCTTACCAAGAGGAGCAATCCACTTTTGG 240
 |||
 Db 208 ATGACGTTACCAATCTGTAGTGTGGCTTACCAAGAGGAGCAATCCACTTTTGG 267
 |||
 Oy 241 GCGAGCGACAGAGCGCGGAAAGCATCTCTTACCGGACACCAAGATCCTCTCTC 300
 |||
 Db 268 GCGAGCGACAGAGCGCGGAAAGCATCTCTTACCGGACACCAAGATCCTCTCTC 327
 |||
 Oy 301 CCATTCACCGAAGCTACAGATCTGAGCGATACCGCGTCAATAGGAGCAAGATCCCT 360
 |||
 Db 328 CCATTCACCGAAGCTACAGATCTGAGCGATACCGCGGATAGGAGCAAGATCCCT 387
 |||
 Oy 361 CTGGATATAGAGAACTCATTCATCCGTCCGCGCTTGCTTATCCAGCGGAGCAGC 420
 |||
 Db 388 CTGGATATAGAGAACTCATTCATCCGTCCGCGCTTGCTTATCCAGCGGAGCAGC 447
 |||
 Oy 421 GCGGCCACAGCTGTTCCCTTAATCTCTCATTCAGATGATCTCGAGGCGCGAGATTC 480
 |||
 Db 448 CGTACCCAGCTGTTCCCTTAATCTCTCATTCAGATGATCTCGAGGCGCGAGATTC 507
 |||
 Oy 481 AATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGGGGAGATCATTTCTCCGAC 540
 |||
 Db 508 AATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGGGGAGATCATTTCTCCGAC 567
 |||
 Oy 541 ATGTACATCTGAGCTGAGACTAGTTGGGCGCCAAATCCAGCAAGTCCAGCACTCT 600
 |||
 Db 568 ACGTACATCTGAGCTGAGACTAGTTGGGCGCCAAATCCAGCAAGTCCAGCACTCT 627
 |||
 Oy 601 ACGATAGCGGTTTTTAATACCATTTTCGTTGGTATATTCACCGGATCACTTCTGACG 660
 |||
 Db 628 ACGATAGCGGTTTTTAATACCATTTTCGTTGGTATATTCACCGGATCACTTCTGACG 687
 |||
 Oy 661 TTGAGCAATGTTGCGAGCTGATGCCAGCTTACGATGATGTTGTTATGAGGAC 720
 |||
 Db 688 TTGAGCAATGTTGCGAGCTGATGCCAGCTTACGATGATGTTGTTATGAGGAC 747
 |||
 Oy 721 CGACATCTTCTCTC 734
 |||
 Db 748 GCGCATCTTCTCTC 761
 |||

RESULT 15
 LOCUS AX019439 762 bp DNA linear PAT 07-SEP-2000
 DEFINITION Sequence 16 from Patent WO9940109.
 ACCESSION AX019439
 VERSION AX019439.1 GI:10043378
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 artificial construct.
 REFERENCES
 1 (bases 1 to 762)
 Weilers,P., Stiefel,T., Voelter,W. and Morris,P.
 Recombinant mistletoe lectins
 Patent: WO 9940109-A 16 12-AUG-1999;
 WEILERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH
 (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB)
 location/Qualifiers
 1..762
 /organism="synthetic construct"

Thu Mar 27 07:06:55 2003

us-09-627-165e-15.rge

Page 11

BASE COUNT 171 a /db_xref="taxon:32630"
ORIGIN 230 c 183 g 178 t

Query Match 75.5%; Score 631.6; DB 6; Length 762;
Best Local Similarity 91.3%; Pred. No. 1.3e-178;
Matches 670; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

```
QY 1 CATCAGACGACGGGCGAGATATATTCGGTTTCATCAGCTTCGACGATATATGTCTCA 60
   ||||| 11 ||||| 1111 ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28 CATCAACCCAGGGGAGGAGATCTCCGGTTTCATCAGCTTCGACGATATATGTCTCA 87
QY 61 AGCGAGAGCTTTTCCATAGATACGACTGCTGCTGCTACGATCCCGCTCGGAT 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88 AGCGAGAGCTTTTCCATAGATACGACTGCTGCTGCTACGATCCCGCTCGGAT 147
QY 121 GCGCAAGATTTGTGTGTGTGAATCCCAATCAGGGGGAGACTGATCAGCGCCGC 180
   ||||| ||||| 111111 111111 111111 111111 111111 111111
Db 148 GCGCAAGATTTGTGTGTGTGAATCCCAATCAGGGGGAGACTGATCAGCGCCGC 207
QY 181 ATCGAGCTTACCAATCTGTAGCTGTGCTTACCAAGCAGCGCAATCTACTTTTG 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 208 ATCGAGCTTACCAATCTGTAGCTGTGCTTACCAAGCAGCGCAATCTACTTTTG 267
QY 241 CCGGAGCAGCAGCGGCGGAAAGCAATCTTTCACCGGACACCAATCTCTCTCTC 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 268 CCGGAGCAGCAGCGGCGGAAAGCAATCTTTCACCGGACACCAATCTCTCTCTC 327
QY 301 CCATTACCGGAGCTACACAGATCTGTAGCGATACCGCGCTCATAGGACAGATCCCT 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 CCATTACCGGAGCTACACAGATCTGTAGCGATACCGCGCTCATAGGACAGATCCCT 387
QY 361 CTGGGTATAGAGAACTCATTCATTCGCTGCGCGCTTCGTATCCAGGGGCGACACC 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 388 CTGGGTATAGAGAACTCATTCATTCGCTGCGCGCTTCGTATCCAGGGGCGACACC 447
QY 421 CCGGCGCAAGCTGCTCCCTTATATCTCATTCAGATGATCTCGAGCGCGAGATTC 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 448 CGTACCCAGAGCTCGTTGATTTATTCATTCAGATGATCTCGAGCGCGAGATTC 507
QY 481 AATCCCATCTTTTGGAGGGCTCGCAATATTAACAGCGGGAGTCAATTTCTCCGAC 540
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 508 AATCCCATCTTTTGGAGGGCTCGCAATATTAACAGTGGGGCTCATTTCTGCGAC 567
QY 541 ATGTACATCTCTGAGCTGTGAGACTAGTTGGGGCAACAATCCAGAGTCCAGACTCT 600
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 568 GTGTACATCTGTGAGCTGTGAGACTAGTGGGGCAACAATCCAGAGTCCAGACTCT 627
QY 601 ACGGATGCGCTTTTAAATTAACCATTTTCGTTGGGTATATCCACCGGTAATCTGTGAG 660
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 628 ACCGATGCGCTTTTAAATTAACCATTCGTTGGGTATATCCCGGTAATCTGTGAG 687
QY 661 TTGAGCAATGTTCCGAGAGTGTATGAGCTTACGATATGTTGTATGTAGAGGAG 720
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 688 TTGAGCAATGTTCCGAGAGTGTATGAGCTTACGATATGTTGTATGTAGAGAG 747
QY 721 CGACCATCTTCTCTC 734
   ||||| ||||| |||||
Db 748 GGGCATCTTCTCTC 761
   ||||| ||||| |||||
```

Search completed: March 26, 2003, 19:07:06
Job time : 204.32 secs

PA	(VISC-) VISCUM AG.
XX	
XX	
PI	Kleff S;
XX	
DR	WPI: 2002-316737/36.
DR	P-PSDB: ABB79450.
XX	
PT	New nucleic acid encoding preprotein of mistletoe lectin, useful as
PT	diagnostic and therapeutic agents, also encodes polypeptide -
XX	
PS	Claim 1; Fig 1; 6pp; German.
XX	
CC	The invention relates to a nucleic acid molecule (ABL56947) that encodes
CC	a preprotein (ABB79450) which, after maturation, has the biological
CC	activity of the galactose-recognising mistletoe lectin (MLIII). The MLIII
CC	encoding nucleic acid molecule, primers specific to it or complements of
CC	it, and encoded (oligomeric) polypeptides are useful as diagnostic and
CC	therapeutic agents.
XX	
XX	
SQ	Sequence 1656 BP; 387 A; 418 C; 449 G; 402 T; 0 other;

Query Match	94.78;	Score 792.8;	DB 24;	Length 1656;
Best Local Similarity	96.88;	Pred. No. 8,9e-249;		
Matches 809; Conservative	0;	Mismatches 27;	Indels 0;	Gaps 0

[illegible]

QY	721	CGACCATCTCTCCGACGCTGCGTATTGGCCGCTGGTATACGACCCGCTTTGGAAAT	780
QY <td>721</td> <td>CGACCATCTCTCCGACGCTGCGTATTGGCCGCTGGTATACGACCCGCTTTGGAAAT</td> <td>780</td>	721	CGACCATCTCTCCGACGCTGCGTATTGGCCGCTGGTATACGACCCGCTTTGGAAAT	780
Db	847	CGCCCATCTTCTCCGAGTGGCTATTGGCCGCTGGTATACGACCCGCTTTGGAAAT	906
QY	781	AGCGGCGCGCTCGACGATTAAGTGCACGTCGTCGCAACCCACGCTCCGATGAT	836
Db	907	AGCGGCGCGCTCGACGATTAAGTGCACGTCGTCGCAACCTACCGCTGGGATGTT	962

RESULT 2
AAT91659
ID AAT91659 standard; DNA; 1923 BP.

AC AAT91659;

18-DEC-1997 (first entry)

DE Prepro mistletoe lectin

Mistletoe; lectin; cytotoxic; A chain; B chain; dimer; ss

OS Viscum album.

FH	Key	Location/Qualifiers
EE	1740	

/*tag= a
304
FT

```
/*tag= b
/*tag= "Mr. gene fragment h"
```

misc_feature	205..909
--------------	----------

```

/note= "ML gene fragment 1
010 057

```

```
/*tag= d  
"MI. gene fragment e"  
/oto-
```

misc_feature	958..1746
FT	4+30= 8

/note= "ML gene fragment 9
1747 1023
FT

```
/*tag= f
/*tag= "Mr. gene fragment i
```

XX 2

XX:

XX
XX
0000 01000019

XX 0570-0100049

XX

XX

XX

DR P-PSDB; 10021.

aa Nucleic acid encoding pre:pro form of mist
PT

XX
F1
XX

XX
F3
XX

It can be used in immunotoxins therapy.

CC acid ligandness. The
CC AAT91659) comprises an A chain (AAT91660)
CC

Sequence 1923 BP; 465 A; 488 C; 495 G; 47

Query Match	Score
83.18;	695.2;

Best Local Similarity	0.9999	0.9999
Matches	758	Conservative
	0	Mismatch

QY 1 CATCAGACGAGCGGCGACGAATATTCGGTTCAATCAGCGCTTCGAGATTAATGTCGA 60
 DB 181 CATCAAAACGAGGCGGAGGAGATATTCGGTTCAATCAGCGCTTCGAGATTAATGTCGA 240
 QY 61 AGCGGAAGCTTTTCCATGAGATACCACTCTTGGCTAGTCTACAGATCCCGTCTCGAT 120
 DB 241 AGCGGAAGCTTTTCCATGAGATACCACTCTTGGCTAGTCTACAGATCCCGTCTCGAT 120
 QY 121 GCGGAAGATTTTGTGTGTGACTACCAATAGGGGGAGACTTCATAGGGCGCC 180
 DB 301 GCGGAAGATTTTGTGTGTGACTACCAATAGGGGGAGACTTCATAGGGCGCC 180
 QY 181 ATGAGGATTTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 240
 DB 361 ATGAGGATTTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 240
 QY 241 GCGGAGCGACGAGCGGCGGAGAAAGCAATCTTACCGGCGGAGCAACGAGATCTCTTC 300
 DB 421 GCGGAGCGACGAGCGGCGGAGAAAGCAATCTTACCGGCGGAGCAACGAGATCTCTTC 300
 QY 301 CCATTACCGGAGGATACAGATCTGAGCGGATACCGGCTATAGGAGCAAGATCCCT 360
 DB 481 CCATTACCGGAGGATACAGATCTGAGCGGATACCGGCTATAGGAGCAAGATCCCT 360
 QY 361 CTGGGTATAGAGAACTCATTCATCCGCTCTGCGCTTCCGTTATCCAGCGGCGAGAC 420
 DB 541 CTGGGTATAGAGAACTCATTCATCCGCTCTGCGCTTCCGTTATCCAGCGGCGAGAC 420
 QY 421 CGGGCCCAAGCTCTTCCCTTATATCTCATTCAGATGATCTCCGAGCGCGAGATTTC 480
 DB 601 CGTACCCAGGCTCTGCTGATTTTAACTTCAATCCGTCAGATGATCTCCGAGCGCGAGATTTC 480
 QY 481 AATCCATCTTTTGGAGGCTCTCCATATCAATTAACGCGGAGGAGCTATTTCTCCGAG 540
 DB 661 AATCCATCTTTTGGAGGCTCTCCATATCAATTAACGCGGAGGAGCTATTTCTCCGAG 540
 QY 541 ATGATCAATCTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 600
 DB 721 ATGATCAATCTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 600
 QY 601 ACCGATGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
 DB 781 ACCGATGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
 QY 721 GAGCATCTCTTCCGAGGAGCTGATTTGCGGCGGCTGATTTGCGGCGGCTGATTTGCGG 780
 DB 901 GAGCATCTCTTCCGAGGAGCTGATTTGCGGCGGCTGATTTGCGGCGGCTGATTTGCGG 780
 QY 901 GAGCATCTCTTCCGAGGAGCTGATTTGCGGCGGCTGATTTGCGGCGGCTGATTTGCGG 780
 DB 954 -----AGCGATGATGTTACTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1001
 QY 781 AGCGGCGGCTGAGGATGTTACTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 836
 DB 954 -----AGCGATGATGTTACTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1001
 RESULT 3
 ID AAV74182 standard: DNA, 1923 BP.
 AC AAV74182:
 XX 30-Apr-1999 (first entry)
 XX Mistletoe lectin prepro-protein DNA.
 XX ML: mistletoe lectin; MC: transgenic plant; glycosylation;
 KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;
 XX cancer; ss.
 XX Viscum album.
 OS

FH Key Location/Qualifiers
 FT CDS 55..1749
 FT /tag- a
 FT /product- "mistletoe lectin"
 EP84388-AL.
 DB 16-DEC-1998.
 DB 26-JUN-1995; 95EP-0109949.
 DB 26-JUN-1995; 95EP-0109949.
 DB 26-JUN-1995; 98EP-0105660.
 DB (MADU) MADUS KOEHL AG.
 DB Baur A, Eck J, Lentzen H, Zinke H;
 DB MPI: 1999-02582/03.
 DB P-FSDB; AAW90127.
 PT New transgenic plant expressing mistletoe lectin - useful for
 PT producing recombinant lectin in e.g. cancer diagnosis and therapy
 XX Claim 1a; Fig 4c; 30pp; German.
 XX This invention describes a novel transgenic plant transformed with a
 XX vector capable of encoding a mistletoe (Viscum album) lectin
 XX preproprotein or a biologically active fragment. The specific
 XX also describes a polypeptide produced by a plant where the polypeptide
 XX exhibits at least one enzymatic modification other than the glycosylation
 XX that occurs in Viscum album or the polypeptide is a fusion protein, a
 XX mistletoe lectin polypeptide dimer and an immunotoxin comprising the
 XX polypeptide or the polypeptide dimer. The plants are used for large-scale
 XX production of mistletoe lectin for diagnostic or therapeutic purposes
 XX (e.g. in cancer therapy). This sequence encodes the mistletoe lectin
 XX used in the method of the invention.
 SQ Sequence 1923 BP; 465 A; 488 C; 495 G; 475 T; 0 other:
 Query Match 83.1%; Score 695.2; DB 20; Length 1923;
 Best Local Similarity 90.7%; Pred. No. 8-5e-217;
 Matches 758; Conservative 0; Mismatches 63; Indels 15; Gaps 1;
 QY 1 CATCAGACGAGCGGCGACGAATATTCGGTTCAATCAGCGCTTCGAGATTAATGTCGA 60
 DB 181 CATCAAAACGAGGCGGAGGAGATATTCGGTTCAATCAGCGCTTCGAGATTAATGTCGA 240
 QY 61 AGCGGAAGCTTTTCCATGAGATACCACTCTTGGCTAGTCTACAGATCCCGTCTCGAT 120
 DB 241 AGCGGAAGCTTTTCCATGAGATACCACTCTTGGCTAGTCTACAGATCCCGTCTCGAT 120
 QY 121 GCGGAAGATTTTGTGTGTGACTACCAATAGGGGGAGACTTCATAGGGCGCC 180
 DB 301 GCGGAAGATTTTGTGTGTGACTACCAATAGGGGGAGACTTCATAGGGCGCC 180
 QY 181 ATGAGGATTTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 240
 DB 361 ATGAGGATTTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 240
 QY 241 GCGGAGCGACGAGCGGCGGAGAAAGCAATCTTACCGGCGGAGCAACGAGATCTCTTC 300
 DB 421 GCGGAGCGACGAGCGGCGGAGAAAGCAATCTTACCGGCGGAGCAACGAGATCTCTTC 300
 QY 301 CCATTACCGGAGGATACAGATCTGAGCGGATACCGGCTATAGGAGCAAGATCCCT 360
 DB 481 CCATTACCGGAGGATACAGATCTGAGCGGATACCGGCTATAGGAGCAAGATCCCT 360
 QY 361 CTGGGTATAGAGAACTCATTCATCCGCTCTGCGCTTCCGTTATCCAGCGGCGAGAC 420
 DB 541 CTGGGTATAGAGAACTCATTCATCCGCTCTGCGCTTCCGTTATCCAGCGGCGAGAC 420
 QY 421 CGGGCCCAAGCTCTTCCCTTATATCTCATTCAGATGATCTCCGAGCGCGAGATTTC 480

```

Db 601 CGTACCAAGCTCGTCGATTTTAATCCATCATGATGATTCGGAGGCCGCAATTC 660
Qy 481 AATCCCATCTTTGGAGGCTCCCAATATCAATTAAGAGCGGAGATTTCTCCGAG 540
Db 661 AATCCCATCTTTGGAGGCTCCCAATATCAATTAAGAGCGGAGATTTCTCCGAG 720
Qy 541 ATGTACATGCTCGAGCTGAGACTAGTGGGCCCAATATCAATTAAGAGCGGAG 600
Db 721 GTGTACATGCTCGAGCTGAGAGCAATTTGGGCCCAATATCAATTAAGAGCGG 780
Qy 601 ACGATGAGCGTTTAAATTAACCCATTTTCGTTGGGTAATATCAATTAAGAGCG 660
Db 781 ACGATGAGCGTTTAAATTAACCCATTTTCGTTGGGTAATATCAATTAAGAGCG 840
Qy 661 TTGAGCAATGTTTCGAGAGCTGATCGCACTTAAGCAATATTTGTTGATGAGGAG 720
Db 841 TTGAGCAATGTTTCGAGAGCTGATCGCACTTAAGCAATATTTGTTGATGAGGAG 900
Qy 721 CGACATCTTCCTCCGAGCTGAGCTAATTTGGCCGCTGCTATACAGACCGCTTG 780
Db 901 CGACATCTTCCTCCGAGCTGAGCTAATTTGGCCGCTGCTATACAGACCGCTTG 953
Qy 781 AGCGGCGCGTTCGAGATGTTTACCTGACACTGCTTCGCAACCCAGCGTGCATCT 836
Db 954 -----AGCGGATGATGTTTACCTGACACTGCTTCGCAACCTAGGATGAT 1001

RESULT 4
AAZ09103 standard: DNA: 1596 BP.
AAZ09103:
18-OCT-1999 (first entry)
Mistletoe lectin I DNA fragment.
Mistletoe lectin: antitumor; immunostimulant; A-chain; MIA; immunity;
ribosome 28S subunit; non-cytotoxic; T-cell activation; immune response;
lymphokine-producing macrophage; uncontrolled cell growth; treatment;
cancer; cytotoxicity; antigen; isoform; lectin I; ds.
Viscum album.
DE19804210-A1.
12-AUG-1999.
03-FEB-1998: 98DE-1004210.
03-FEB-1998: 98DE-1004210.
(BIOS-) BIOSYN ARZNEIMITTEL GMBH.
Morris P, Stiefel T, Voelter W, Welters P:
WPI: 1999-44535/38.
P-PSDB: AAY25979.
Preparation of mistletoe lectins in heterologous systems,
particularly for use as anticancer agents and immunostimulants
Claim 14: Fig 1A: 78pp: German.
This invention describes a novel mistletoe lectin (I) and its fragments
which have antitumor and immunostimulatory activity. The A-chain (MIA)
of the mistletoe lectin binds to, and inactivates, the 28S subunit of
ribosomes. Non-cytotoxic forms of (I) activate T-cell and
lymphokine-producing macrophages, so stimulate immunity. (I) and its
fragments are used to treat uncontrolled cell growth (particularly
cancers) and if they lack cytotoxicity to increase the strength of the
immune response, particularly to a co-administered antigen

```

```

Cc (tumour-associated, bacterial or viral). The method allows production of
Cc mistletoe lectin, and its individual chains, in many different isoforms
Cc and on a large scale, at any time of the year. Recombinant products are
Cc free from toxins present in natural mistletoe extracts. This sequence
Cc encodes a fragment of the mistletoe lectin I protein.
S0 Sequence 1596 BP: 387 A; 418 C; 419 G; 372 T; 0 other:
Query Match 82.7%; Score 692; DB 20; Length 1596;
Best Local Similarity 90.4%; Pred. No. 8.7e-216;
Matches 756; Conservative 0; Mismatches 65; Indels 15; Gaps 1;
Qy 1 CATCAGACGAGGCGGAGCAATATTTCCGCTTCATCAGCTTCTCCGAGATATGATCA 60
Db 28 CATCAACACGAGGCGGAGCAATATTTCCGCTTCATCAGCTTCTCCGAGATATGATCA 87
Qy 61 AGCGGAGCTTTTCCAAATGAGATACCACTTTCGCTGATAGATCCCGTCTCGAT 120
Db 88 AGCGGAGCTTTTCCAAATGAGATACCACTTTCGCTGATAGATCCCGTCTCGAT 147
Qy 121 GCGCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 180
Db 148 GCGCAAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 207
Qy 181 ATGAGATTTACCAATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 240
Db 208 ATGAGATTTACCAATGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 267
Qy 241 GCGGAGCGACAGAGCGGCGGAGGAGGATCTCTTCACGCGGACACAGATCTCTTC 300
Db 268 GCGGAGCGACAGAGCGGCGGAGGAGGATCTCTTCACGCGGACACAGATCTCTTC 327
Qy 301 CCATTACCGGACACTTACAGATCTGAGAGGATAGCGGCTGATAGGAGGAGGAGCT 360
Db 328 CCATTACCGGACACTTACAGATCTGAGAGGATAGCGGCTGATAGGAGGAGGAGCT 387
Qy 361 CTGGGATAGAGCAATCATCATTCATTCGCTGCGGCTGCTATACAGAGGAGGAG 420
Db 388 CTGGGATAGAGCAATCATCATTCATTCGCTGCGGCTGCTATACAGAGGAGGAG 447
Qy 421 CGGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 448 CGGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 507
Qy 481 AATCCCATCTTTTGGAGGCTCGCCATATATTAACAGCGGAGGAGTCTATTTCCG 540
Db 508 AATCCCATCTTTTGGAGGCTCGCCATATATTAACAGCGGAGGAGTCTATTTCCG 567
Qy 541 ATGTACATGCTCGAGCTGAGAGCTAGTTGGGCGCAATATCAATTAAGAGCGG 600
Db 568 GTGTACATGCTCGAGCTGAGAGCTAGTTGGGCGCAATATCAATTAAGAGCGG 627
Qy 601 ACGATGAGCGTTTAAATTAACCCATTTTCGTTGGGTAATATCAATTAAGAGCG 660
Db 628 ACGATGAGCGTTTAAATTAACCCATTTTCGTTGGGTAATATCAATTAAGAGCG 687
Qy 661 TTGAGCAATGTTTCGAGAGCTGATCGCACTTAAGCAATATTTGTTGATGAGGAG 720
Db 688 TTGAGCAATGTTTCGAGAGCTGATCGCACTTAAGCAATATTTGTTGATGAGGAG 747
Qy 721 CGACATCTTCCTCCGAGCTGAGCTAATTTGGCCGCTGCTATACAGACCGCTTG 780
Db 748 CGACATCTTCCTCCGAGCTGAGCTAATTTGGCCGCTGCTATACAGACCGCTTG 800
Qy 781 AGCGGCGCGTTCGAGATGTTTACCTGACACTGCTTCGCAACCCAGCGTGCATCT 836
Db 801 -----AGCGGATGATGTTTACCTGACACTGCTTCGCAACCTAGGATGAT 1001

RESULT 5
AAZ09100 standard: DNA: 1598

```

AC AA209100;
XX
XX 18-OCT-1999 (first entry)
XX
XX Mistletoe lectin DNA consensus sequence.
XX
XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
KM riboxyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
KM lymphokine-producing macrophage; uncontrolled cell growth; treatment;
KM cancer; cytotoxicity; antigen; Isoform; ss.
XX
XX Viscum album.
OS
XX DE19804210-A1.
PN
XX 12-ANG-1999.
PD
XX 03-FEB-1998; 98DE-1004210.
PF
XX 03-FEB-1998; 98DE-1004210.
PR
XX 03-FEB-1998; 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
PA
XX Morris P, Stiefel T, Voelter W, Welters P;
PI
XX WPI; 1999-445335/38.
DR
XX
XX Preparation of mistletoe lectins in heterologous systems,
PT particularly for use as anticancer agents and immunostimulants
XX
XX Claim 11; Page 32-33; 78pp; German.
XX
XX This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC represents a consensus sequence of mistletoe lectin DNA described in the
CC specification.
XX
XX
SQ Sequence 1598 BP; 373 A; 396 C; 392 G; 356 T; 81 other;
Query Match 81.7%; Score 683.8; DB 20; Length 1598;
Best Local Similarity 87.2%; Pred. No. 4.2e-213;
Matches 730; Conservative 37; Mismatches 54; Indels 16; Gaps 2;
OY 1 CATCAGACGAGCGGCGACGAATATTCGCGTCATCAGCGCTTCGAGATATATGCTCA 60
DB 28 CATCAACACGACGCGGCGAKGAACTACTCCGGTCATCAGCGCTTCGAGATATATGCTCA 87
OY 61 AGCGGAAGCTTTTCCATATGATACCACTCTTGGCTGACGTACGATCCCGCTCGGAT 120
DB 88 AGCGGAAGCTTTTCCATATGATACCACTCTTGGCTGACGTACGATCCCGCTCGGAT 147
OY 121 GCGCAAGATTTTGTGTGAGAACTACCAATCAGGCGGAGACATCGATCAGGCGGCC 180
DB 148 GCGCAAGATTTTGTGTGAGAACTACCAATCAGGCGGAGACATCGATCAGGCGGCC 207
OY 181 ATCGACGTTACCAATCTGTACGTGTGGCTTACCAAGCAGCGACCAATCTCTTTTG 240
DB 208 ATCGACGTTACCAATCTGTACGTGTGGCTTACCAAGCAGCGACCAATCTCTTTTG 267
OY 241 GCGCAGCAGCAGCGCGGCGGAAAGCATCTTCCACCGGACACACAGCA-TCTCTCTT 299
DB 268 GCGCAGCAGCAGCGCGGCGGAAAGCATCTTCCACCGGACACACAGCAATCTCTCTT 327

OY 300 CCCATTACCGGAGAGCTACAGATCTGAGCGATACCGCGGCTCATAGGACAGATCCC 359
DB 328 CCCATTACCGGAGAGCTACAGATCTGAGCGATACCGCGGCTCATAGGACAGATCCC 387
OY 360 TCTGGGTATAGAGAACTCATTCATCCGCTCGGCGCTTGTATCCAGCGGACGAC 419
DB 388 TCTCGGTATAGAGAACTCATTCATCCGCTCGGCGCTTGTATCCAGCGGACGAC 447
OY 420 CCGGCGGCGGAGCGTCTCCCTATATACCTCATTCAGATGATGATCCGCGGCGGAGATT 479
DB 448 GCGTTCATAGCTCGTTGATTTATCTCTCATTCATTCAGATGATGATCCGCGGCGGAGATT 507
OY 480 CATCCCATCTTTTGGAGGCGCTCGGCAATACATTAACAGCGGAGATCATTTCTCCGA 539
DB 508 CATCCCATCTTTATGAG 567
OY 540 CATGTACATGCTCGAGCTGAGAGACTAGTTGGGCGCCACATTCACGCAAGTCCAGAGTC 599
DB 568 CRTGTACATGCTCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
OY 600 TACGAGTGGCGGCTTTTATATACCCATTTCCGTTGGGTATATCCACCGGTAAGTCTGAGAC 659
DB 628 AACCGATGGCGTTTATATATACCCATTTCCGTTGGGTATATCCACCGGTAAGTCTGAGAC 687
OY 660 GTTGAGCAATGTTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTTATGAGGGA 719
DB 688 GTTGAGCAATGTTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTTATGAGGGA 747
OY 720 CCGACCATCTTCTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTTATGAGGGA 779
DB 748 GCGGCGCATCTTCTCCGAGAGCTGATCGCCAGCTTACGATCATGTTGTTTATGAGGGA 801
OY 780 TAGCGGCGCGCTGAGAGCTGATCGCCAGCTTACGATCATGTTGTTTATGAGGGA 836
DB 802 -----AGCGGATGATGTTACTGCAAGTCTGAGAGCTGAGAGCTGAGGATTTG 849
RESULT 6
AAC85472 standard; cDNA; 762 BP.
ID AAC85472;
XX AAC85472;
AC 16-MAY-2001 (first entry)
XX
XX
XX
XX A-chain gene isoform for biosynthesis of a Korean mistletoe lectin #1.
XX
XX Isoform; A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KM KML; tumour; KM-110; KML-C; KMLBP; KML-ITU; KML-ILU;
KM heparin binding protein; ds.
XX
XX Viscum album coloratum.
OS
XX EP1074560-A2.
PN
XX 07-FEB-2001.
PD
XX 27-JUL-2000; 2000EP-0402158.
PF
XX 27-JUL-1999; 99KR-0030638.
PR
XX
XX (MIST-) MISTLE BIOTECH CO LTD.
PA
XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI Kang T, Park C;
PI
XX WPI; 2001-171044/18.
DR
XX P-PSDB; AAB47090.
XX
XX Novel lectin proteins isolated from Korean mistletoe, useful for
PT enhancing immunity and effectuating anti-tumoral activity -
XX
XX Claim 1; Page 25; 62pp; English.

XX The sequences given in AAC85472-74 encode isoforms of an A-chain gene
 CC involved in biosynthesis of lectins isolated from Korean mistletoe.
 CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
 CC for treating tumours. The KML's are isolated from a protein fraction
 CC derived from the leaves, stems and fruits of Korean mistletoe, which
 CC is designated KM-110. One of the isolates, KM-C was shown to be
 CC effective against colon 26-M3.1 carcinoma and L5178Y-M25 lymphoma.
 XX

Sequence 762 BP; 172 A; 220 C; 186 G; 184 T; 0 other;

Query Match 79.9%; Score 668.4; DB 22; Length 762;
 Best Local Similarity 94.4%; Pred. No. 3.3e-208;
 Matches 693; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```

OY 1 CATCAGACGACGGGCGAGCAATATTTCCGGTTCATCAGCGCTTCCGAGATTATGTCTCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28 CATCAAAACCACGGGCGAGCAATATTTCCGGTTCATCAGCGCTTCCGAGATTATGTCTCA 87
OY 61 AGCGGAGCTTTTCCATGAGATACACTCTTGCCTGCTAGCATCCCGCTCGGAT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88 AGCGGAGCTTTTCCATGAGATACACTCTTGCCTGCTAGCATCCCGCTCGGAT 147
OY 121 GCGCAAGATTTGTGTGTGGAACTCAACAATCAGGGGGAGACTCGATCAGCGCCGC 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 148 GCGCAAGATTTGTGTGTGGAACTCAACAATCAGGGGGAGACTCGATCAGCGCCGC 207
OY 181 ATCGAGTTACCATCTAGCTGTGGCTTACCAAGAGGCGAGCCATCTACTTTTG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 208 ATCGAGTTACTTACCTTACGTGTGGCTTACCAAGAGGCGAGCAATCTACTTTTG 267
OY 241 GCGGACGACACGACGCGCGGAAAGGCAATCTTTCACCGGACACACAGATCCTCTC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 268 GCGGACGACACGACGCGCGGAAAGGCAATCTTTCACCGGACACACAGATCCTCTC 327
OY 301 CCATTCACCGGAGGTATACAGATGTGAGAGATACCGCGCTCATAGGACAGATCCCT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 CCATTCACCGGAGGTATACAGATGTGAGAGATACCGCGCTCATAGGACAGATCCCT 387
OY 361 CTGGTATAGAGAGACTCATTCATCCGTCTCGGCGCTTCTGTTATCCAGGGGAGCAC 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 388 CTGGGTATAGAGAGACTCATTCATCCGTCTCGGCGCTTCTGTTATCCAGGGGAGCAC 447
OY 421 GCGGCGCAAGCTCGTCCCTTATTAATCTCATTCAGATGATCTCCGAGCGCGAGATTC 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 448 CTGGCGCAAGCTCGTCCCTTATTCATCTCATTCAGATGATCTCCGAGCGCGAGATTC 507
OY 481 AATCCATCTTTGAGGGGCTCGGCAATACATTAACAGCGGGGAGTCAATTTCTCCGAC 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 508 AATCCATCTTTGAGGGGCTCGGCAATACATTAACAGCGGGGAGTCAATTTCTCCGAC 567
OY 541 ATGTACATGCTGAGCTGAGACTAGTTGGGGCAACAAATCACGCAAGTCCAGAGTCT 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 568 ACGTACATTTCTGAGCTGAGAGCGAGTTGGGGCAACAAATCACGCAAGTCCAGAGTCT 627
OY 601 ACGGATGGCGTTTATTAATACCAATTCGTTGGGTATATCCACCGGTAACTTCGTGAC 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 628 ACGGATGGCGTTTATTAATACCAATTCGTTGGGTATATCCACCGGTAACTTCGTGAC 687
OY 661 TTGACCAATGTTGCGAGAGTATGCGCACTTAGGATCATGTTGTATGAGAGAC 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 688 TTGACCAATGTTGCGAGAGTATGCGCACTTAGGATCATGTTGTATGAGAGAC 747
OY 721 CGACCATCTTCCTC 734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 CGGCAATCTTCCTC 761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 7
 AAT91660
 ID AAT91660 standard; DNA; 774 BP.
 XX
 AC AAT91660;

XX
 DT 18-DEC-1997 (first entry)
 XX
 DE Prepro mistletoe lectin A chain.
 XX
 KW Mistletoe; lectin; cytotoxic; A chain; B chain; dimer; ss.
 XX
 OS Viscum album.
 XX

Key Location/Qualifiers
 FT CDS 4..765
 FT /*tag= a

EP751221-AL.

02-JAN-1997.

26-JUN-1995; 95EP-0109949.

26-JUN-1995; 95EP-0109949.

(MADU) MADANUS KOELIN AG.

Baur A, Eck J, Lentzen H, Zinke H;

WPI: 1997-054678/06.

P-PSDB; AAM10022.

Nucleic acid encoding pre:pro form of mistletoe lectin - for

therapeutic or diagnostic use

Claim 2; Fig 4A; 30pp; German.

Mistletoe lectin is a cytotoxic agent that has been used for tumour

therapy. It can be used in immunotoxins and medicaments. Nucleic

acid fragments can be used in diagnostic methods. Mistletoe lectin (

AAT91659) comprises an A chain (AAT91660) and a B chain (AAT91661).

Sequence 774 BP; 173 A; 232 C; 185 G; 184 T; 0 other;

Query Match 77.1%; Score 645.2; DB 18; Length 774;
 Best Local Similarity 92.7%; Pred. No. 1.3e-200;
 Matches 677; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

```

OY 1 CATCAGACGACGGGCGAGCAATATTTCCGGTTCATCAGCGCTTCCGAGATTATGTCTCA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 34 CACGACACGACCGGTGAGATATTTCCGGTTCATCAGCGCTTCCGAGATTATGTCTCA 93
OY 61 AGCGGAAGCTTTTCCATGAGATACCACTTTCGCTCAGTCTACGATCCCGTCTCGAT 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 94 AGCGGAAGCTTTTCCATGAGATACCACTTTCGCTCAGTCTACGATCCCGTCTCGAT 153
OY 121 GCGCAAGATTTGTGTGTGGAACTCAACAATCAGGGGGAGACTCGATCAGCGCCGC 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 GCGCAAGATTTGTGTGTGGAACTCAACAATCAGGGGGAGACTCGATCAGCGCCGC 213
OY 181 ATCGAGTTACCAATCTGATGAGTGTGGCTTACCAAGCAGGAGCAATCTACTTTTG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 214 ATCGAGTTACCAATCTGATGAGTGTGGCTTACCAAGCAGGAGCAATCTACTTTTG 273
OY 241 GCGGACGACACGAGCGCGGAAAGGCAATCTCTTACCGGACCAACAGATCTCTCTC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 274 GCGGACGACACGAGCGCGGAAAGGCAATCTCTTACCGGACCAACAGATCTCTCTC 333
OY 301 CCATTCACGGAAGCTTACAGATCTGAGAGGATGCGCGGTATAGGGATCCAGATTCCT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 334 CCATTCACGGAAGCTTACAGATCTGAGAGGATGCGCGGTATAGGGATCCAGATTCCT 393
OY 361 CTGGTATAGAGAGACTCATTCATCCGTCTGCGGCTTCTTATTCAGGCGGACAGACC 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 394 CTGGTATAGAGAGACTCATTCATCCGTCTGCGGCTTCTTATTCAGGCGGACAGACC 453
OY 421 GCGGCGCAAGCTCGTCCCTTATTAATCTCATTCAGATGATCTCCGAGGCGCGAGATTC 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the

XX

Mistletoe rMLA variant DNA.

PT Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 PS Claim 12; Page 34; 78pp; German.
 XX
 CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumor and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of mistletoe lectin A-chain (MLA) DNA
 CC described in the specification.
 XX
 XX Sequence 763 BP; 163 A; 215 C; 175 G; 170 T; 40 other;
 SQ
 Query Match 76.68; Score 641; DB 20; Length 763;
 Best Local Similarity 89.18; Pred. No. 3.1e-199;
 Matches 655; Conservative 37; Mismatches 42; Indels 1; Gaps 1;
 QY 1 CATGACGACGAGGCGGCGAATATTCGGTTCATCAGCGTCTCCGAGATTATGTCTCA 60
 DB 28 CATCAACGACGCGGCGGCGAATATTCGGTTCATCAGCGTCTCCGAGATTATGTCTCA 87
 QY 61 AGCGGAAGCTTTTCCAAATGAGATACCACTCTGCTAGTCTAGTATCCCGCTCTCGAT 120
 DB 88 AGCGGAAGCTTTTCCAAATGAGATACCACTCTGCTAGTCTAGTATCCCGCTCTCGAT 147
 QY 121 GCGCAAGATTTGTGTGTGGAACTCACCACATCAGGCGGGAACATCTGATCAGCGCGCC 180
 DB 148 GCGCAAGATTTGTGTGTGGAACTCACCACATCAGGCGGGAACATCTGATCAGCGCGCC 207
 QY 181 ATGACGTTACCAATCTGATGCTGTGGCTTACCAAGCAGGCGCAATCTCTTTTG 240
 DB 208 ATGACGTTACCAATCTGATGCTGTGGCTTACCAAGCAGGCGCAATCTCTTTTG 267
 QY 241 GCGGAGCGACGACGCGGCGGAAAGGATCTCTTACCGGCGCACACAGA-TCTCTCTG 299
 DB 268 GCGGAGCGACGACGCGGCGGAAAGGATCTCTTACCGGCGCACACAGANTCTCTCT 327
 QY 300 CCATTCACGCGAAGCTACACAGATGTGGAGCATACGCGGCTCATAGGAGCAGATCCC 359
 DB 328 CCATTCACGCGAAGCTACACAGATGTGGAGCATACGCGGCTCATAGGAGCAGATCCC 387
 QY 360 TCTGGTATAGAGAACTCATTCATCCGCTCTGGCGCTTCTGTTATCCAGGCGGACGAC 419
 DB 388 TCTGGTATAGAGAACTCATTCATCCGCTCTGGCGCTTCTGTTATCCAGGCGGACGAC 447
 QY 420 CCGGCGCGAAGCTGTTCCCTTAATCTCATATTCAGATATTCGAGCGCGGAGATT 479
 DB 448 GCGTTCACGAGCTGCTGCTGATTTATCTCATTCAGATATTCGAGCGCGGAGATT 507
 QY 480 CAATCCCATTTTGGGGGCTGCGCAATACATTACACGCGGGAATCTTTCTCCCA 539
 DB 508 CAATCCCATTTTGGGGGCTGCGCAATACATTACACGCGGGAATCTTTCTCCCA 567
 QY 540 CATGTACATGCTGAGCTGAGAGCTAGTTGGGCGCAACATCCAGCAGATCCAGAGC 599
 DB 568 CAGTACATGCTGAGCTGAGAGCTAGTTGGGCGCAACATCCAGCAGATCCAGAGC 627
 QY 600 TAGGAGATGGGTTTAAATTAACCATTTGGGTTATATTCACCGGTAATCTCTGAC 659
 DB 628 AACCGATGGGTTTAAATTAACCATTTGGGTTATATTCACCGGTAATCTCTGAC 687
 QY 660 GTTGACCAATGTTGCGAGTATGCGGACGATAGGAGCATGTTTGTGTATGTGGGA 719
 DB 688 GTTGACCAATGTTGCGAGTATGCGGACGATAGGAGCATGTTTGTGTATGTGGGA 747

QY 720 CCAGCATCTCTCTC 734
 DB 748 GCGGCGATCTCTCTC 762
 RESULT 11
 AAV51341
 ID AAV51341 standard; DNA; 762 BP.
 AC AAV51341;
 AC 23-OCT-1998 (first entry)
 DE Mistletoe rMLA DNA.
 XX
 KW Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; ricin; translocation; ss.
 XX
 OS Viscum album.
 XX
 XX Key Location/Qualifiers
 FT CDS 4..759
 FT /*tag= a
 FT /product= "Lectin A-chain"
 FT /note= "Partial sequence"
 XX
 PN W09829540-A2.
 XX
 PD 09-JUL-1998.
 XX
 PF 02-JAN-1998; 98WO-EP00009.
 XX
 PR 02-JAN-1997; 97EP-0100012.
 XX
 PA (BRAI-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
 XX
 PI Eck J, Schmidt A, Zinke H;
 DR WPI; 1998-388122/33.
 DR P-PSDB; AAW64659.
 XX
 PT Nucleic acid encoding fusion protein containing mistletoe lectin A
 PT chain - useful for treatment of proliferative and autoimmune
 PT diseases, allergies and tumours
 PS Disclosure; Fig 11a; 115pp; German.
 XX
 CC This sequence encodes a lectin A-chain, rMLA, isolated from mistletoe.
 CC This sequence can be used in the construction of a fusion protein which
 CC comprises an effector module that is cytotoxic intracellularly, a
 CC processing module covalently bonded to the effector module and
 CC containing a protease recognition sequence, and a targeting module
 CC covalently bonded to the processing module, able to bind specifically to
 CC the surface of a cell so as to mediate internalisation of the fusion
 CC protein. Such a fusion protein can be used for treating disorders
 CC involving proliferation and/or elevated activation of cells, especially
 CC autoimmune disease, allergy and tumours. The proteins can be administered
 CC e.g. by injection or topically but especially by intravenous injection,
 CC at 1 ng to 500 mg g/kg/day, or for ex vivo use at 1 pg to 500 ng/ml.
 CC Fusion proteins can develop toxic activity in a wide range of target
 CC cells. The processing module prevents extracellular dissociation, and
 CC fusion proteins based on mistletoe lectin A-chain are far more active
 CC than those based on ricin and do have the associated problems of
 CC non-specific toxicity. The protein may be expressed in a non-glycosylated
 CC form that does not bind to sugar receptors in the liver, and which has a
 CC long half-life in the blood. Where the mistletoe lectin B-chain is used,
 CC it actively assists in translocation of the ML A-chain from the
 CC endoplasmic reticulum to the cytoplasm.
 XX
 SQ Sequence 762 BP; 169 A; 230 C; 183 G; 180 T; 0 other;


```

Db      568  ACGTACATTCCTCAGAGCTGGAGACGAGTTGGGGGCAACATCCAGCAAGTTCACAGACTCG 627
Qy      601  ACGATGGCGCTTTTAAATAAACCATTTGGTGGTATATCCACCGTAACTTCGTGACG 660
Db      628  ACGGATGGCGCTTTTAAATAAACCATTTGGTGGTATATCCACCGTAACTTCGTGACG 687
Qy      661  TTGAGCAATGTTGGCGAGCTGATCGCGCACTTAGCATCATGTTGTTGATGTAGGAC 720
Db      688  TTGAGCAATGTTGGCGAGCTGATCGCGCACTTAGCATCATGTTGTTGATGTAGGAC 747
Qy      721  CGACCATCTTCCTC 734
Db      748  CGGCATCTTCCTC 761

```

RESULT 13

AAZ09104

ID AAZ09104 standard; DNA; 762 BP.

```

AC      AAZ09104;
XX      18-OCT-1999 (first entry)
DE      Mistletoe lectin A1 DNA fragment.

```

Mistletoe lectin; antitumour; immunostimulant; A-chain; MLA; immunity; ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response; lymphokine-producing macrophage; uncontrolled cell growth; treatment; cancer; cytotoxicity; antigen; isoform; lectin A1; ds.

Viscum album.

DE19804210-A1.

PD 12-AUG-1999.

XX 03-FEB-1998; 98DE-1004210.

XX 03-FEB-1998; 98DE-1004210.

XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

XX Morris P, Stiefel T, Voelter W, Welters P;

XX MPI; 1999-445335/38.

XX P-PSDB; AAY25980.

Preparation of mistletoe lectins in heterologous systems, particularly for use as anticancer agents and immunostimulants

XX Claim 15; Fig 2A; 78pp; German.

This invention describes a novel mistletoe lectin (I) and its fragments which have antitumour and immunostimulatory activity. The A-chain (MLA) of the mistletoe lectin binds to, and inactivates, the 28S subunit of ribosomes. Non-cytotoxic forms of (I) activate T-cell and lymphokine-producing macrophages, so stimulate immunity. (I) and its fragments are used to treat uncontrolled cell growth (particularly in cancers) and if they lack cytotoxicity, to increase the strength of the immune response, particularly to a co-administered antigen (tumour-associated, bacterial or viral). The method allows production of mistletoe lectin, and its individual chains, in many different isoforms and on a large scale, at any time of the year. Recombinant products are free from toxins present in natural mistletoe extracts. This sequence encodes a fragment of the mistletoe lectin A1 protein.

XX Sequence 762 BP; 171 A; 230 C; 183 G; 178 T; 0 other;

Query Match

75.5%; Score 631.6; DB 20; Length 762;

Best Local Similarity 91.3%; Pred. No. 3.7e-196;

Matches 670; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

```

Qy      1  CATCAGACGAGGGGCGACGAATATTCGGTTATACGCTTCCGAGATTATGTCGA 60
Db      28  CATCAAAACCGAGGGCGGAGAACTTCGGTTATACGCTTCCGAGATTATGTCGA 87
Qy      61  AGCGAAGCTTTTCCATGATGATACCACTTCGTCAGTCAGTACGATGCCGTCGGAT 120
Db      88  AGCGAAGCTTTTCCATGATGATACCACTTCGTCAGTCAGTACGATGCCGTCGGAT 147
Qy      121  GCGCAAGATTTGTGGTGTGAACCTACCAATCAGGGGAGACTGATCAAGCGCGC 180
Db      148  GCGCAAGATTTGTGGTGTGAACCTACCAATCAGGGGAGACTGATCAAGCGCGC 207
Qy      181  ATGAGGTTACCAATCTGATCGTGGCTTACCAAGCAGGGACCAATCTATTTTG 240
Db      208  ATGAGGTTACCAATCTGATCGTGGCTTACCAAGCAGGGACCAATCTATTTTG 267
Qy      241  CGGAGCAGCAGGAGCGGAGAAAGCAATCTTCACGGGACGACCAATCTCTCTC 300
Db      268  CGGAGCAGCAGGAGCGGAGAAAGCAATCTTCACGGGACGACCAATCTCTCTC 327
Qy      301  CCAATCAGCGGAGAGCTACACAGATCTGAGCGATACCGGTCATAGGAGCAGATCCCT 360
Db      328  CCAATCAGCGGAGAGCTACACAGATCTGAGCGATACCGGTCATAGGAGCAGATCCCT 387
Qy      361  CTGGGATAGAGAACTCATTCATCCGTCGCGCTTCGTTATCCAGCGGACGACCC 420
Db      388  CTGGGATAGAGAACTCATTCATCCGTCGCGCTTCGTTATCCAGCGGACGACCC 447
Qy      421  CGGAGCAGCAGTCCCTTATATCCATTCAGATGATCCGAGGCGCGGAGATTG 480
Db      448  CGTACCCAGAGCTGTCATTTAATCCATTCAGATGATCCGAGGCGCGGAGATTG 507
Qy      481  AATCCATCTTTTGAAGGCTCGCCATATACATTAACAGGGGAGTCAATTTCTCCGAC 540
Db      508  AATCCATCTTTTGAAGGCTCGCCATATACATTAACAGGGGAGTCAATTTCTCCGAC 567
Qy      541  ATGTACATGCTGGAGCTGAGACTAGTTGGGGCGAAATCCAGCAAGTCCAGAGCT 600
Db      568  GTGTACATGCTGGAGCTGAGACTAGTTGGGGCGAAATCCAGCAAGTCCAGAGCT 627
Qy      601  ACGGATGGCGCTTTTAAATAAACCATTTCCGTTGGGATATCCACCGTAATCTCGTACG 660
Db      628  ACGGATGGCGCTTTTAAATAAACCATTTCCGTTGGGATATCCACCGTAATCTCGTACG 687
Qy      661  TTGAGCAATGTTGGCGAGCTGATCGCGCACTTAGCATCATGTTGTTGATGTAGGAC 720
Db      688  TTGAGCAATGTTGGCGAGCTGATCGCGCACTTAGCATCATGTTGTTGATGTAGGAC 747
Qy      721  CGACCATCTTCCTC 734
Db      748  CGGCATCTTCCTC 761

```

RESULT 14

AAZ09105

ID AAZ09105 standard; DNA; 768 BP.

XX 18-OCT-1999 (first entry)

DE Mistletoe lectin A2 DNA fragment.

Mistletoe lectin; antitumour; immunostimulant; A-chain; MLA; immunity; ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response; lymphokine-producing macrophage; uncontrolled cell growth; treatment; cancer; cytotoxicity; antigen; isoform; lectin A2; ds.

Viscum album.

DE19804210-A1.

PD 12-AUG-1999.

XX 03-FEB-1998; 98DE-1004210.
PF
XX 03-FEB-1998; 98DE-1004210.
PR
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
PA
XX Morris P, Stiefel T, Voelter W, Welters P;
PI WPI: 1999-445335/38.
DR P-PSDB; AAY25981.
XX
XX Preparation of mistletoe lectins in heterologous systems,
PT particularly for use as anticancer agents and immunostimulants
PS
PS Claim 15; Fig 3A; 78pp; German.
XX
XX This invention describes a novel mistletoe lectin (I) and its fragments
CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
CC fragments are used to treat uncontrolled cell growth (particularly
CC cancers) and if they lack cytotoxicity, to increase the strength of the
CC immune response, particularly to a co-administered antigen
CC (tumour-associated, bacterial or viral). The method allows production of
CC mistletoe lectin, and its individual chains, in many different isoforms
CC and on a large scale, at any time of the year. Recombinant products are
CC free from toxins present in natural mistletoe extracts. This sequence
CC encodes a fragment of the mistletoe lectin A2 protein.
XX
SQ Sequence 768 BP; 175 A; 224 C; 182 G; 187 T; 0 other;
Query Match 73.4%; Score 614; DB 20; Length 768;
Best Local Similarity 90.4%; Pred. No. 2.1e-190;
Matches 669; Conservative 0; Mismatches 65; Indels 6; Gaps 1;
OY 1 CATCAGACGAGCGGCGAGCAATATTTCGGTTCATCAGCTTCCGAGATTATGTCGA 60
DB 28 CATCAACACAGCGGGGAGTAATCTCCGGTTCATCAGCTTCCGAGATTATGTCGA 87
OY 61 AGCGGAGCTTTTCCATGAGATACACTTTCGGTTCAGTACGATCCCGTCCGAT 120
DB 88 AGCGGAGCTTTTCCATGAGATACACTTTCGGTTCAGTACGATCCCGTCCGAT 147
OY 121 GCGCAAGATTGTGTGGTGAAGTACCAATCAGAGGAGAGTCCGATCAGCGGCC 180
DB 148 GCGCAAGATTGTGTGGTGAAGTACCAATCAGAGGAGAGTCCGATCAGCGGCC 207
OY 181 ATCGAGCTTACCAATCTGTAGGTGGCTTACCAAGCAGGCGCAATCTTCTTTTG 240
DB 208 ATCGAGCTTACCAATCTGTAGGTGGCTTACCAAGCAGGCGCAATCTTCTTTTG 267
OY 241 GCGGAGGAGCAGAGGCGGCGGAGGAGGAGTCTTACCGGAGGAGGAGGAGGAGG 294
DB 268 GCGGAGGAGCAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 327
OY 295 TCTCTCCATTCACCGGAGGTACACAGATCTGAGGAGTACGCGGTCTATAGGAGCCAG 354
DB 328 TCTCTCCATTCACCGGAGGTACACAGATCTGAGGAGTACGCGGTCTATAGGAGCCAG 387
OY 355 ATCCCTGTGGGTATAGAGAACTCATTAATCCGTCTGCGGCTTCTGTTATCCAGGCGGC 414
DB 388 ATCCCTGTGGGTATAGAGAACTCATTAATCCGTCTGCGGCTTCTGTTATCCAGGCGGC 447
OY 415 AGCAGCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 474
DB 448 AGCAGCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 507
OY 475 AGATTCAATCCCATCTTTTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 534
DB 508 AGATTCAATCCCATCTTTTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 567

OY 535 CCCGACATGTACATGCTGAGCTGAGACTAGTTGGGCGCAATCCACGCAAGTCCAG 594
DB 568 CCAGACATGTACATGCTGAGCTGAGAGCTAGTTGGGCGCAATCCACGCAAGTCCAG 627
OY 595 CAGTCTACGAGATGGCTTTTATATACCATTTGGTGGTATATCCACCGGTACATTC 654
DB 628 CATTCACCGATGGCGTTTATATACCATTTGGTGGTATATCCACCGGTACATTC 687
OY 655 GTGACGTTGAGCAATGTTGGCGGAGTATCGCAGCTTACGATCATGTTTGTATAGT 714
DB 688 GTGACGTTGCTATATGTTGCTCTGTGATCGCAGCTTGGCATCATGTTTGTATAGC 747
OY 715 AGGAGCGACCATCTTCTTC 734
DB 748 GGAGAGCGGCGCATCTTCTTC 767
RESULT 15
AAC85474
ID AAC85474 standard; cDNA; 768 BP.
XX
AC AAC85474;
XX
XX 16-MAY-2001 (first entry)
DT
XX
XX A-chain gene isoform for biosynthesis of a Korean mistletoe lectin #3.
DE
XX
XX Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KW KML; tumour; KM-110; KML-C; KMBP; KML-IIU; KML-III;
KW heparin binding protein; ds.
XX
XX Viscum album coloratum.
OS
XX
PN EP1074560-A2.
PD
XX
XX 07-FEB-2001.
XX
XX 27-JUL-2000; 2000BP-0402168.
PF
XX
XX 27-JUL-1999; 99KR-0030638.
PR
XX
PA (MIST-) MISTLE BIOTECH CO LTD.
XX
XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI Kang T, Park C;
PI
XX
DR WPI: 2001-171044/18.
DR
XX
XX P-PSDB; AAB47092.
PT
XX
XX Novel lectin proteins isolated from Korean mistletoe, useful for
PS enhancing immunity and effectuating anti-tumoral activity -
PS
PS Claim 5; Page 27-28; 62pp; English.
XX
XX The sequences given in AAC85472-74 encode isoforms of an A-chain gene
CC involved in biosynthesis of lectins isolated from Korean mistletoe.
CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
CC for treating tumours. The KML's are isolated from a protein fraction
CC derived from the leaves, stems and fruits of Korean mistletoe, which
CC is designated KM-110. One of the isolates, KML-C was shown to be
CC effective against colon 26-M3.1 carcinoma and L5178Y-ML25 lymphoma.
XX
XX
SQ Sequence 768 BP; 171 A; 219 C; 200 G; 178 T; 0 other;
Query Match 70.3%; Score 588.6; DB 22; Length 768;
Best Local Similarity 89.1%; Pred. No. 4.4e-182;
Matches 662; Conservative 0; Mismatches 69; Indels 12; Gaps 2;
OY 1 CATCAGACGAGCGGCGAGCAATATTTCGGTTCATCAGCTTCCGAGATTATGTCGA 60
DB 28 CATCAACACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 87
OY 61 AGCGGAGCTTTTCCATGAGATACCACTTTCGCTCAGCTACGATCCCGGCTTCGAT 120

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 18:18:22 ; Search time 50.4724 Seconds

(without alignments)

5085.723 Million cell updates/sec

Title: US-09-627-165E-15

Sequence: 837

1 catcagacgacggcgacga.....aacccacgtgcacgtca 837

Scoring table:

IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCrUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	695.2	83.1	1923	4 US-08-776-059-34	Sequence 34, App1
2	645.2	77.1	774	4 US-08-776-059-30	Sequence 30, App1
3	628.2	75.1	705	4 US-08-776-059-38	Sequence 38, App1
4	78.8	9.4	1140	1 US-08-218-303-15	Sequence 15, App1
5	78.8	9.4	1140	2 US-08-338-793D-60	Sequence 60, App1
6	77.2	9.2	807	2 US-08-356-786-7	Sequence 7, App1
7	77.2	9.2	1605	2 US-08-356-786-9	Sequence 9, App1
8	77.2	9.2	1855	4 US-09-147-208-33	Sequence 33, App1
9	77.2	9.2	1855	4 US-09-147-208-40	Sequence 40, App1
10	77.2	9.2	1855	4 US-09-147-208-47	Sequence 47, App1
11	77.2	9.2	1855	4 US-09-147-208-54	Sequence 54, App1
12	77.2	9.2	1879	4 US-09-147-208-23	Sequence 23, App1
13	77.2	9.2	1879	4 US-09-147-208-25	Sequence 24, App1
14	77.2	9.2	1879	4 US-09-147-208-25	Sequence 25, App1
15	66.4	7.9	813	1 US-08-425-336-11	Sequence 11, App1
16	66.4	7.9	813	1 US-08-425-336-11	Sequence 11, App1
17	66.4	7.9	813	1 US-08-425-336-11	Sequence 11, App1
18	66.4	7.9	813	1 US-08-425-336-11	Sequence 11, App1
19	66.4	7.9	813	2 US-08-646-360-11	Sequence 11, App1
20	66.4	7.9	813	2 US-08-621-803-246	Sequence 246, App1
21	66.4	7.9	813	3 US-08-839-765-11	Sequence 11, App1
22	66.4	7.9	813	3 US-09-136-389-11	Sequence 11, App1
23	66.4	7.9	813	4 US-09-217-352-246	Sequence 246, App1
24	66.4	7.9	813	4 US-09-217-352-246	Sequence 246, App1
25	65	7.9	774	5 PCT-US91-05766-1	Sequence 11, App1
26	65	7.8	955	4 US-08-621-803-258	Sequence 258, App1
27	65	7.8	1003	2 US-08-621-803-252	Sequence 252, App1

28	65	7.8	1003	4 US-08-217-352-252	Sequence 252, App1
29	65	7.8	1072	2 US-08-621-803-250	Sequence 250, App1
30	65	7.8	1072	2 US-09-217-352-250	Sequence 250, App1
31	63.2	7.6	813	1 US-07-901-707-11	Sequence 11, App1
32	63.2	7.6	813	1 US-07-901-707-11	Sequence 11, App1
33	63.2	7.6	813	1 US-07-988-430-57	Sequence 57, App1
34	63.2	7.6	813	1 US-07-988-430-57	Sequence 57, App1
35	63.2	7.6	813	5 PCT-US92-09487-11	Sequence 11, App1
36	63.2	7.6	813	5 PCT-US92-09487-11	Sequence 11, App1
37	59	7.0	913	1 US-07-901-707-13	Sequence 13, App1
38	59	7.0	913	1 US-07-988-430-13	Sequence 13, App1
39	59	7.0	913	1 US-08-425-336-13	Sequence 13, App1
40	59	7.0	913	1 US-08-425-336-13	Sequence 13, App1
41	59	7.0	913	1 US-08-425-336-13	Sequence 13, App1
42	59	7.0	913	1 US-08-425-336-13	Sequence 13, App1
43	59	7.0	913	2 US-08-646-360-13	Sequence 13, App1
44	59	7.0	913	3 US-08-839-765-13	Sequence 13, App1
45	59	7.0	913	3 US-09-136-389-13	Sequence 13, App1
				4 US-09-610-838-13	Sequence 13, App1

ALIGNMENTS

```
RESULT 1
US-08-776-059-34
Sequence 34, Application US/08776059B
Patent No. 6271368
GENERAL INFORMATION:
APPLICANT: LENTZEN, Hans
APPLICANT: ECK, Jurgen
APPLICANT: BAUR, Axel
APPLICANT: ZINKE, Holger
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
FILE REFERENCE: 674503-2003
CURRENT APPLICATION NUMBER: US/08/776,059B
EARLIER FILING DATE: 1999-06-19
EARLIER APPLICATION NUMBER: PCT/EP96/02273
EARLIER FILING DATE: 1996-06-25
EARLIER APPLICATION NUMBER: 9510949.8
EARLIER FILING DATE: 1995-06-26
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 34
LENGTH: 1923
TYPE: DNA
ORGANISM: Viscum album
US-08-776-059-34

Query Match
Best Local Similarity 83.1%; Score 695.2; DB 4; Length 1923;
Matches 758; Conservative 0; Mismatches 63; Indels 15; Gaps 1;

QY 1 CATCAGACGACGGCGAGCATATTTCCGCTATCAGCGCTCCGAGATTATGCTCA 60
    |||||
DB 181 CATCAACACGACGGCGAGCATATTTCCGCTATCAGCGCTCCGAGATTATGCTCA 240
QY 61 AGCGAAGCTTTTCATGATGATATCCATCTTGGCTAGTCTACGATCCCGCTCGGAT 120
    |||||
DB 241 AGCGAAGCTTTTCATGATGATATCCATCTTGGCTAGTCTACGATCCCGCTCGGAT 300
QY 121 GCGCAAGATTGTTGTTGATGAGCAATCACCACAGGGGAGACATCATCAGCGCGC 180
    |||||
DB 301 GCGCAAGATTGTTGTTGATGAGCAATCACCACAGGGGAGACATCATCAGCGCGC 360
QY 181 ATCAGACTTACCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
    |||||
DB 361 ATCAGACTTACCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 241 CCGGACGACCAACGACGCGGAGGAGGATCTTACCGGACGACCAACGATCTCTCTC 300
    |||||
DB 421 CCGGACGACCAACGACGCGGAGGAGGATCTTACCGGACGACCAACGATCTCTCTC 480
QY 301 CCATTTCACCGAAGCTTACACAGATCTGAGAGCATGCGCGGTCTATGAGGACCATCTCCT 360
```

Query Match	77.1%;	Score 645.2;	DB 4;	Length 774;
Best local similarity	92.7%;	Pred. No. 1.2e-207;		
Matches 677;	Conservative	0;	Mismatches 53;	Indels 0; Gaps:
QY	1	CATCAGACGACGGCGACGATATTTTCGGTCAACACAGCTTCCCGAGATTATGTC	60	
Db	34	CACCGACACACCGGTGAAGATATTTTCGGTCAACAGCTTCCCGAGATTATGTC	93	
QY	61	AACGGAACTTTTCATGATGATACCACTTCGCTGAGTCAGATCCCGTCTCGAT	12	
Db	94	AGGGGAAGCTTTTCATGATGATACCACTTCGCTGAGTCAGATCCCGTCTCGAT	15	

Query Match	75.1%	Score 628.2;	DB 4;	Length 705;
Best local Similarity	93.2%;	Pred. No. 6.4e-202;		
Matches 657; Conservative	0;	Mismatches 48;	Indels 0;	Gaps 0

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 08:25:55 ; Search time 6.57789 Seconds
(without alignments)
833.848 Million cell updates/sec

Title: US-09-627-165e-14
Perfect score: 596
Sequence: 1 ARFNPIMRLRQINSSESS.....VVDANDVTCTXSEPTVATV 119

Scoring table: PROSUM62
Gapop 10.0, Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications, AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCIT_NEM_PUB pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCUTS_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*
10: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*
12: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	311	52.2	252	10	US-09-347-064-2
2	311	52.2	252	10	US-09-347-064-8
3	772	12.1	247	9	US-09-792-793A-39
4	70.5	11.8	293	10	US-09-765-527-259
5	70.5	11.8	332	10	US-09-765-527-251
6	69.5	11.7	309	10	US-09-765-527-253
7	69.5	11.7	477	9	US-09-738-626-645
8	69	11.6	247	9	US-09-792-793A-34
9	68.5	11.5	251	10	US-09-765-527-247
10	68.5	11.5	1018	10	US-09-801-574-32
11	65.5	11.0	937	10	US-09-949-197-19
12	65	10.9	1110	9	US-10-149-819-12
13	64.5	10.8	951	9	US-09-924-097-15
14	64	10.7	263	10	US-09-347-064-10
15	64	10.7	267	10	US-09-347-064-4
16	62.5	10.5	136	9	US-10-020-419-4
17	62	10.4	366	10	US-09-764-864-808
18	61	10.2	1229	10	US-09-815-242-5818
19	61	10.2	1229	10	US-09-815-242-12946

20	60.5	10.2	260	10	US-09-742-693-31	Sequence 31, Appl
21	60	10.1	115	10	US-09-746-801A-47	Sequence 47, Appl
22	60	10.1	144	10	US-09-864-761-37049	Sequence 37049, A
23	60	10.1	211	10	US-09-728-911-34	Sequence 34, Appl
24	60	10.1	212	9	US-09-912-672A-6	Sequence 6, Appl
25	60	10.1	240	10	US-09-864-761-47947	Sequence 38293, A
26	60	10.1	312	10	US-09-864-761-38293	Sequence 5, Appl
27	60	10.1	560	9	US-09-912-672A-5	Sequence 164, Appl
28	60	10.1	574	9	US-09-912-672A-2	Sequence 164, Appl
29	60	10.1	574	9	US-10-063-547-164	Sequence 164, App
30	60	10.1	574	9	US-10-063-616-164	Sequence 164, App
31	60	10.1	574	9	US-10-063-502-164	Sequence 25, Appl
32	60	10.1	574	10	US-09-728-911-25	Sequence 4, Appl
33	60	10.1	574	10	US-09-870-574-4	Sequence 164, App
34	60	10.1	574	12	US-10-006-867-164	Sequence 33522, A
35	60	10.1	777	10	US-09-864-761-33522	Sequence 2, Appl
36	59.5	10.0	552	10	US-09-835-654-2	Sequence 522, App
37	59	9.9	110	10	US-09-925-302-522	Sequence 4, Appl
38	59	9.9	262	10	US-09-978-274A-4	Sequence 11800, A
39	59	9.9	263	10	US-09-978-274A-4	Sequence 2, Appl
40	59	9.9	268	10	US-09-815-242-11800	Sequence 35, Appl
41	59	9.9	314	10	US-09-978-274A-2	Sequence 44, Appl
42	58.5	9.8	275	9	US-09-801-368-44	Sequence 14, Appl
43	58.5	9.8	876	10	US-09-738-626-4478	
44	58	9.7	413	9	US-09-479-614-14	
45	58	9.7	431	9	US-09-479-614-14	

ALIGNMENTS

```
RESULT 1
US-09-347-064-2
; Sequence 2, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jürgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-inactivating Proteins of the mistletoe Viscum
; TITLE OF INVENTION: album
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347, 064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Viscum album
US-09-347-064-2

Query Match      52.2%   Score 311:   DB 10:   Length 252:
Best local similarity 71.8%   Pred. No. 2.6e-30;
Matches 61:   Conservative 7;   Mismatches 17;   Indels 0;   Gaps 0;
```

OY 1 ARFNPIMRLRQINSSESSPPNNYMLELTSNGROSTGYOQSKDGFNTQRLQISAGN 60
DB 168 ARFNPIMRLRQINSSESSPPNNYMLELTSNGROSTGYOQSKDGFNTQRLQISAGN 227
OY 61 FVTXSNVADVLSLAIMLFECCGRP 85
DB 228 FVTXSNVADVLSLAIMLFECCGRP 252

RESULT 2
US-09-347-064-8
; Sequence 8, Application US/09347064A

Patent No. US20020045208A1
GENERAL INFORMATION:
APPLICANT: Eck, Jürgen
APPLICANT: Schmidt, Arno
APPLICANT: Zinke, Holger
TITLE OF INVENTION: Recombinant Fusion Proteins Based on
TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
FILE REFERENCE: 09282-5
CURRENT APPLICATION NUMBER: US/09/347,064A
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: PCT/EP98/00009
EARLIER FILING DATE: 1998-01-02
EARLIER APPLICATION NUMBER: EP 97 10 0012.0
EARLIER FILING DATE: 1997-01-02
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 252
TYPE: PRT
ORGANISM: Viscum album
US-09-347-064-8

Query Match 52.2%; Score 311; DB 10; Length 252;
Best Local Similarity 71.8%; Pred. No. 2,6e-30;
Matches 61; Conservative 7; Mismatches 17; Indels 0; Gaps 0;
QY 1 ARPNPFXWRLRQINSGSSPPNMVMELETSWGRSTOVQOSKDGIFNQIRLQISAGN 60
DB 167 ARPNPFXWRLRQINSGSSPPNMVMELETSWGRSTOVQOSKDGIFNQIRLQISAGN 60
QY 61 FVYXSNVRDVISSLAIMLFEGSGRPSLDHPS 85
DB 227 FVLTNVRDVISSLAIMLFEGSGRPSLDHPS 251

RESULT 3
US-09-792-793A-39
Sequence 39, Application US/09792793A
Patent No. US20020168370A1
GENERAL INFORMATION:
APPLICANT: McDonald, John R.
APPLICANT: Coggin, Philip
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAMAGE AND
TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
FILE REFERENCE: 25020-601D
CURRENT APPLICATION NUMBER: US/09/792,793A
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 93
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3;
LENGTH: 247
TYPE: PRT
ORGANISM: Trichosantheus kirilowii
US-09-792-793A-39

Query Match 12.1%; Score 72; DB 9; Length 247;
Best Local Similarity 26.8%; Pred. No. 0.46;
Matches 22; Conservative 24; Mismatches 30; Indels 6; Gaps 3;
QY 1 ARENPFXWRLRQINSGSSPPNMVMELETSWGRSTOVQOSKDGIFNQIRLQISAGN 58
DB 162 ARENPFXWRLRQINSGSSPPNMVMELETSWGRSTOVQOSKDGIFNQIRLQISAGN 58
QY 59 GNVYXSNVRDVISSLAIMLFEGSGRPSLDHPS 78
DB 220 GNVYXSNVRDVISSLAIMLFEGSGRPSLDHPS 219

RESULT 4
US-09-765-527-251
Sequence 251, Application US/09765527
Patent No. US20020006638A1

GENERAL INFORMATION:
APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Roole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 259:
SEQUENCE CHARACTERISTICS:
LENGTH: 293 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 259:
US-09-765-527-259

Query Match 11.8%; Score 70.5; DB 10; Length 293;
Best Local Similarity 26.6%; Pred. No. 0.89;
Matches 25; Conservative 19; Mismatches 47; Indels 3; Gaps 2;
QY 1 ARENPFXWRLRQINSGSSPPNMVMELETSWGRSTOVQOSKDGIFNQIRLQISAGN 59
DB 190 ARENPFXWRLRQINSGSSPPNMVMELETSWGRSTOVQOSKDGIFNQIRLQISAGN 59
QY 60 NEVYXSNVRDVISSLAIMLFEGSGRPSLDHPS 93
DB 248 NEVYXSNVRDVISSLAIMLFEGSGRPSLDHPS 281

RESULT 5
US-09-765-527-251
Sequence 251, Application US/09765527
Patent No. US20020006638A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Roole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

Qy	25	TTCCGTTTCATCACGCTTCTCCGAGATTATGTCTCAAGCGGAAGCTTTTCCAAATGAGATA	84
Db	1	TTCCGGTTTCATCACGCTTCTCCGAGATTATGTCTCAAGCGGAAGCTTTTCCAAATGAGATA	60
Qy	85	CCACTCTTGGCGTCAAGTCTAGATCCCGGTCTCGGATGCGCAAGATTGTGTGGTGAA	1444
Db	61	CCACTCTTGGCGTCAAGTCTAGATCCCGGTCTCCCATCGCAAGATTGTGTGGTGAG	120
Qy	145	CTCAACCAATTCAGGGGGGAGATCTCATACGGCGCGCAATTCAGTGTACCANTGTACTGTG	204
Db	121	CTCAACCAACCAAGGGGGGAGACTCATACAGCGCGCGCAATTCAGTGTACCANTGTACTGTG	180
Qy	205	GTCGCTTACCAAGCAGGCGCAATCTCACTTTTGGCGCAGCAGCAGAGCGGCGGAA	264
Db	181	GTCGCTTACCAAGCAGGCGCAATCTCACTTTTGGCGCAGCAGCAGAGCGGCGGAA	240
Qy	265	AGGATCTCTTCACCGCGACACCAAGATCTCTCTCCCATTCACCGAAGCTACACAGAT	324
Db	241	AGGATCTCTTCACCGCGACACCAAGATCTCTCTCCCATTCACCGAAGCTACACAGAT	300
Qy	325	CTGGAGCGATACGCGCGGTCATATAGGAGCAGATCCCTCTGGGTATATAGAGAACTCATTC	384
Db	301	CTGGAGCGATACGCGCGGTCATATAGGAGCAGATCCCTCTCGGTATATAGCAACTCATTC	360
Qy	385	TTCGCTCTCGCGCTTTCGTTATCAAGGGGCGAGCACCCGCGGCGCAAGCTCGTCCCTATA	444
Db	361	TTCGCTCTCGCGCTTTCGTTATCCCGGGGCGAGCACCGGCTACCAAGCTCGTCCGATTTTA	420
Qy	445	ATCCCATTCATGATGATCTCCGAGCGCGCGAGATTCAATCCCATCTTTTGGAGGCTCCG	504
Db	421	ATCCCATTCATGATGATCTCCGAGCGCGCGAGATTCAATCCCATCTTTATGAGAGGCTCCG	480
Qy	505	CAATACATTAAACACGCGGGGAGTCATTTCTTCCGACATGTATACATAGCTCGAGCTGAGACT	564
Db	481	CAATACATTAAACAGTGGGGGCTCATTTCTGTGCAGACGTGTATACATGCTGGAGCTGGAGAGC	540
Qy	565	AGTTGGGGCCCAACATCCACGCAAGTCCAGCAGTCTAAGSAGGCGTTTAAATTAACCA	624
Db	541	AGTTGGGGCCCAACATCCACGCAAGTCCAGCAGTCTAAGSAGGCGTTTAAATTAACCA	600
Qy	625	TTTCGGTTGGATATATCCACGCGGTAATCTGTGTGACGTTGAGCAATGTTCGCGAGGATC	684
Db	601	ATTGGGTTGGATATATCCCGCGGTAATCTGTGTGACGTTGAGCAATGTTCGCGAGGATC	660
Qy	685	GCCAGCTTAAAGCATGTGTTTGTATGTATGATGAGGAGCCAGCATCT	729
Db	661	GCCAGCTTAAAGCATGTGTTTGTATGTATGATGAGGAGCCAGCATCT	705
RESULT 4			
US-08-218-303-15			
Sequence 15, Application US/08218303			
Patent No. 5547867			
GENERAL INFORMATION:			
APPLICANT: Kara, Bhupendra V.			
APPLICANT: Hockney, Robert C.			
APPLICANT: Pitcon, John E.			
TITLE OF INVENTION: FERMENTATION PROCESS			
NUMBER OF SEQUENCES: 23			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Cushman, Darby & Cushman			
STREET: 1615 L Street, N.W.			
City: Washington			
STATE: D.C.			
COUNTRY: U.S.A.			
ZIP: 20036-5601			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			

```

1  APPLICATION NUMBER:  US/08/218,303
2
3  FILING DATE:
4
5  CLASSIFICATION:  435
6
7  PRIOR APPLICATION DATA:
8
9  APPLICATION NUMBER:  US 07/841,533
10
11 FILING DATE:  26-FEB-1992
12
13 ATTORNEY/AGENT INFORMATION:
14
15 NAME:  Kokulis, Paul N.
16
17 REGISTRATION NUMBER:  16,773
18
19 REFERENCE/DOCKET NUMBER:  PK/3893/94908/MJW
20
21 TELECOMMUNICATION INFORMATION:
22
23 TELEPHONE:  202-861-300C
24
25 TELEFAX:  202-822-0944
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008

```



```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..807
; OTHER INFORMATION: /note= "product = "Ricin-A chain"
US-08-356-786-7

Query Match
Best Local Similarity 49.1%; Pred. No. 3.8e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

QY 124 CAAGATTGTGTGGTGGAGAACTACCAATCAGGGGGAGACATCGATCAGCCGCCATC 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 CAAGGGTTTATTAGTTGAACCTTCAAAATCAGACAGCTTCTGTATCATTAAGCGCTG 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 GACGTTACCAATCTGTACGAGTGGTGGCTTACCAAGCAGCCACCAATCTCTATTTTGG- 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 GATGTACACCATGATATGTGTAGGTACCGTGTCTGAAATAGCCGATATTTCTTTTCAT 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 243 -----CGACGACCAAGACGGGGGAAAGGATCTCTTACCCGACCCACCAGA 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 CCGTACATCAGGAAGATGCAAGCAATCATCTTTTTCACGTATGTTCAAAATCGA 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 TCCCTCTCCCATTCACCGAGGTACACAGATCTGGAGCGATACGCCGTCAAT--AGG 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 TATACATTCGCTTGGTGAATATATGATAGATGAAACACTTGTGCTGAATCTGAGA 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 349 GACCAGATCCCTCTGGGTATAGAGAACTCATTCATCCGTCGCGGCGCTTCTTA---- 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 GAAATATCGAGTTGGGAATGCGCCACTAGAGAGGCTATCTCAGCGCTTTATTTATAC 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 -----TCCAGCGGCGACACCGGGGCCAAAGCTCGTCCCTTAATATCTCATTCAGATG 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 AGTACTGGTGGCACTGACTTCCAACTGTGGCTGCTCTTTAATTTGATCCCAATG 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 ATCTCGAGGCGCGAGATCAATCCCATCTTTTGGAGGCTGGCCATACATTAACAGC 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 526 ATTTGAGAGAGCAAAATTCATATATTTAGAGGAGAAATGGCGCAAGAAATTAAGTAC 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 GGGGAGTCAATTTCTCCGAGATGATACATGCTCGAGCTGGAGACTAGTTGGGCGCAAA 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 586 AACGGAGATCTGACACAGATCTAGGCTAATTAACACTTGAAGATAGTTGGGAGACTT 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 TCCAGCAAGTCCAGAGCTACAGATGGCGTTTAAATACCATTTCGGTTGGGTATA 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 646 TCCACTGCAATTCAAAGGTCTTAACCAAGAGACCTTCTAGTCAATTCACATGCAAGA 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 TCCACGTAACCTCGAGCTGAGCAATGTCGCGAGCTGATCCGACACTTAAGCATC 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 706 CGTATAGGTTCCAAATTCAGTGTGTACGATGTGAGTATTAATCCTATCATAGCTCTC 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 ATGTTGTTGTATGTAGGACGACCATCTTC 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 ATGGTATATAGTCGCACTCCACCATTCGTC 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-08-356-786-9
; Sequence 9, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Rind, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
```

```
;
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1605 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1605
; OTHER INFORMATION: /note= "product = "G-FIT"
US-08-356-786-9

Query Match
Best Local Similarity 9.2%; Score 77.2; DB 2; Length 1605;
Best Local Similarity 49.1%; Pred. No. 5.4e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

QY 124 CAAGATTGTGTGGTGGAGAACTACCAATCAGGGGGAGACATCGATCAGCCGCCATC 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 CAAGGTTTATTAGTTGAACCTTCAAAATCAGACAGCTTCTGTATCATTAAGCGCTG 231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 GACGTTACCAATCTGTACGAGTGGTGGCTTACCAAGAGCGGACCAATCTCTATTTTGG- 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 GATGTACCAATTCATATGTGTAGGCTACCGCTGGAATAGCGCATATTTCTTTCAAT 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 243 -----GAGCGACAGACGGGGGAAAGGATCTCTTACCGGACCCACCA 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 CCGTACATCAGGAAGATGCAAGCAATCATCTTTTTCAGTATGATCAAAATCGA 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 TCCCTCTCCCATTCACCGAGCTACACAGATCTGGAGCGATACGCCGTCAAT--AGG 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 TATACATTCGCTTGGTGTATATATGATAGACTTGAACAACACTGTGTATATGTGAGA 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 349 GACCAGATCCCTGTGGGTATAGAGAACTCAATCAATCCGCTCGCGCTTCTTA---- 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 GAAATATATCAGATGGGAATGTCCACTAGAGAGGCTATCTCACTTTATTAATTAAC 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 -----TCCAGGGGGGACGACCGGGGCCAAAGCTGTCCTCTATATATCTAATAGATG 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 AGTACTGGTGGCACTGACTTCCAACTTGGCTGCTCTTTAATTAATTTGATCCAAATG 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 ATCTCGAGGCGCGAGATTCATCCATCTTTTGGAGGCTGGCCATATTAACAGC 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 ATTTGAGAGAGCAAGATTCATATATTTAGAGGAGAAATGCGACGCAATTAATGTAGAC 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 GGGGAGTCAATTTCTCCGACATGTACATGCTCGAGCTGGAGACTAGTTGGGCGCAACA 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 AACCGGAGATCTGACCAATCTAGCCGTAATTAACATTTGAAATAGTTGGGAGGACTT 651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 TCCAGCAAGTCCAGAGCTACAGATGGCGTTTAAATACCATTTCGGTTGGGTATA 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 TCCACTGCAATTCAGAGTCTTAACCAAGAGACCTTCTAGTCAATTCACATTCACAAAGA 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1855 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-54

```

```

Query Match          9.2%; Score 77.2; DB 4; Length 1855;
Best Local Similarity 49.1%; Pred. No. 5.8e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

```

```

QY 124 CAAGATTGTTGTTGGTGAATCACTACCAATCAGGGGGAGACTGATACGCGCCATC 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 274 CAACGGTTTATTTAGTTGAATCTCAATATCATGACAGAGCTTCTGTACATTAAGCGGTG 333
QY 184 GACGTACCAATCTGTAGTGGTGGTCTACCAAGCAGCGCAATCTACTATTTTGGC- 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 GATGTCACCAATGCATATGTGTGCTGCTACCGCTGCTGGAATAGCGCATATTTCTTTAT 393
QY 243 -----CGACGACACAGACGCGCGGAAAGCATCTTTCACCGGACACACAGA 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 CTTGACATTCAGAGAGATGAGAACCAATCATCTCTTTTCACTGATGTTCAAAATGCA 453
QY 292 TCCCTCTCCCATTCACCGGAAAGTACACAGATCTGGAGCGATACGCGGTAT---AGG 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 454 TATTAATTCGCTTTGGTGTATTAATGATAGACTTGAACAACTTGCTGTAATCTGAGA 513
QY 349 GACGATCCCTCTGGGTATAGAGAACTCAATCCGCTCGCGCTTCGTTA----- 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 514 GAAATATGAGATTGGAAATGTCTCACTAGAGAGGCTATCTCGCGCTTATTTATAC 573
QY 405 -----TCCAGCGGCGACAGACCGGCGCAAGCTGCTCCCTTAATATCTCATTCAGATG 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 574 ACTACTGGTGGCACTACAGCTTCCACTCTGGCTCGTCTTATATTTGATTCCAATG 633
QY 460 ATCTCGAGGCGCGAGATTCATCCCATCTTTTGGAGGGCTCGGCATATTAAACAGC 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 634 ATTTAGAGACAGCAAGATTCATATATTGAGGAGAAATGCGACAGAAATTAAGTAC 693
QY 520 GGGAGGATTTCTTCCGACATGTACATGCTGAGCTGGAGACTATTTGGGGCCACAA 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 694 AACCGAGATCTGACACAGATCTGAGCGTAATTAACCTTGAGAAATATTTGGGGGAGACTT 753
QY 580 TCCAGCGCAAGTCCAGCAGTCTACGATGCGCTTTTAAATTAACCATTTCCGTTGGGTATA 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 754 TCCACTGCATTCAGAGAGCTTAACCAAGAGAGCTTGTAGTCCAAATTCACATTCGAAAGA 813
QY 640 TCCACCGGTAATCTGTGAGCTGTGAGCAATGTTGCGAGCGATGCGCAGGTTAGCATG 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 814 CCTAATGGTCCAAATTCAGTGTGATGAGTATTAATTAATCCATATACGCTCTC 873
QY 700 ATGTTGTTGTATGAGGACCGACCAATCTTC 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 874 ATGGTGTAATGATGGCGACCTCCACATCTGTC 905

```

```

RESULT 12
US-09-147-208-23
; Sequence 23; Application US/09147208
; Patent No. 633303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERSKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-23

```

```

Query Match          9.2%; Score 77.2; DB 4; Length 1879;
Best Local Similarity 49.1%; Pred. No. 5.8e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

```

```

QY 124 CAAGATTGTTGTTGGTGAATCACTACCAATCAGGGGGAGAGACTGATACGCGCCATC 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 298 CAACGGTTTATTTAGTTGAATCTCAATATCATGACAGAGCTTCTGTACATTAAGCGGTG 357
QY 184 GACGTTACCAATCTGTAGTGGTGGTCTTACCAAGCAGCGCAATCTACTTTTGGC- 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 358 GATGTCACCAATGCATATGTGTGCTGCTACCGCTGGAATAGCGCATATTTCTTTCAAT 417
QY 243 -----CGACGACACAGACGCGGCGGAAAGCATCTTTCACCGGACACACAGA 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 418 CTTGACATTCAGAGAGATGAGAAAGCATCATCTTTTCACTGATGTTCAAAATGCA 477
QY 292 TCCCTCTCCCATTCACCGGAAAGTACACAGATCTGGAGGATACCGCGGTAT---AGG 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 478 TATTAATTCGCTTTGGTGTATTAATGATAGACTTGAACAACTCTGTGTAATCTGAGA 537
QY 349 GACGATTCCTCTGGGTATAGAGAACTATTCATTCGCTTCGCGCGCTTCTGTTA----- 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 538 GAAATATTCAGATGGGAAATGTGTCCACTGAGAGAGGCTATTCACGCTTATTAATTTAC 597
QY 405 -----TCCAGCGGCGACACACCGGCGCAAGCTGTTCCCTTAATTCCTCATTCAGATG 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 598 AGTACTGGTGGCACTACGCTTCCAACTCTGGCTCGTCTTATTAATTTGATTCGAAATG 657
QY 460 ATCTCGAGGCGCGAGATTCATTCACATCTTTTGGAGGGCTGCCAATATCAATTAACAGC 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 658 ATTTCAAGACAGCAAGATTCATATATTTGAGGAGAAATGCGACAGAAATTAAGTAC 717
QY 520 GGGAGTCAATTTCTTCCGACATGTACATGCTGAGCTGAGACTGATTGGGGCCACAA 579

```


Db 718 AACCGAGATCTGCACGAGATCTAGCGTATCTAGCTGTAGATGAGTGGGAGACTT 777
QY 580 TCCACGACAGTCCAGACGATCTACGAGATGGCTTTTATTAACCATTTGGTGGTATA 639
Db 778 TCCACTGCAATTCAGATGCTTAACCAAGGACCTTTCCTGCTCAATTCAGTCAAGA 837
QY 640 TCCACCGGTAACTCTGACGCTTACGCAATGTCGCGACGATGCGCAGCTTAGCATC 699
Db 838 CGTATGCTTCCAAATTCAGTGTCTAGATGTGAGTATATTAATCCATCATAGCTCTC 897
QY 700 ATGTTGTTGATGTAGGAGCCAGCATCTTC 731
Db 898 ATGGTGTATAGATGCGCACCTCCACCATCTGC 929

RESULT 13
US-09-147-208-24
; Sequence 24, Application US/09147208
; Patent No. 633303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Rictin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-24

Query Match 9.2%; Score 77.2; DB 4; Length 1879;
Best Local Similarity 49.1%; Pred. No. 5.8e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

QY 124 CAAGAATTTGTGTGGTGAACCTCACCAATCAGGGGGAGACTGATCAGCGCGCATC 183
Db 298 CAACGGTTTATTAGTTGAACCTCTCAATCATCTCAGAGCTTCTGTATCATTAAGCCCTG 357
QY 184 GACCTTACCAATCTGTACGTGTGGCTTACCAAGCAGCGACCAATCTCTTTTTCGG- 242
Db 358 GATGTACCAATGATGTGTGGCTGCTACCGCTGGAATAGCCGATATTTCTTTCAT 417
QY 243 -----CGACGACACAGACGGCGGAAAGGCATCTCTACCGGACACACAGCA 291
Db 418 CCTGACATCAGGAAGATGCAAGAACCAATCATCATCTTTTTCACGTAGTTCAAAATCGA 477
QY 292 TCTCTCTCCCATTCACCGGAAGCTACAGATCTGAGGAGATACGCCGGTCAAT--AGG 348

Db 478 TATACATTCGCCCTTTTGTGGTATATATGATAGACTGAAACACTTCTGTAATCTGAGA 537
QY 349 GACCAATCCCTCTGGGTATAGAGAACTCAATTCATCCGCTCTGGCGCTTGCTTA---- 404
Db 538 GAAATATGAGATGGGAAATGGTCCATAGAGAGAGCTATCTACAGCGCTTATATATTAC 597
QY 405 -----TCCAGGGCGGACAGACCGGGCCCAAGCTCCGTTCCCTATATATCTCATTTGAGATG 459
Db 598 AGTACTGTGGGACCTAGCTTCCAACTCTGGCTCGTTCTTATATATTTGCAATCCAAATG 657
QY 460 ATCTCCGAGCCCGAGATTCATCCCATCTTTTGGAGGGCTGCCAATACATTAACAC 519
Db 658 ATTTCAAGACAGCAAGATTTCCATATATTAGAGGAGAAATCGCACGAGATTTAGTTC 717
QY 520 GGGAGCTATTTCTCCGACATGTACATGCTCGACCTGAGACATAGTTGGGGCCAAACA 579
Db 718 AACCGAGATCTGCACGAGATCTACGATGGCGTTTATTAACCATTTGGTGGTATA 639
QY 580 TCCACGACAGTCCAGAGTCTACGATGGCGTTTATTAACCATTTGGTGGTATA 639
Db 778 TCCACTGCAATTCAGATGCTTAACCAAGAGCGCTTGTAGTCAATTCAGTCAAGA 837
QY 640 TCCACCGGTAACTCTGACGCTTACGCAATGTTCCGCGACGTGATCGCCAGTTAGCATC 699
Db 838 CGTATGCTTCCAAATTCAGTGTGTAGCATGTGAGTATATTAATCCATCATAGCTCTC 897
QY 700 ATGTTGTTGATGTAGGAGCCAGCATCTTC 731
Db 898 ATGGTGTATAGATGCGCACCTCCACCATCTGC 929

RESULT 14
US-09-147-208-25
; Sequence 25, Application US/09147208
; Patent No. 633303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Rictin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-25

Query Match 9.2%; Score 77.2; DB 4; Length 1879;

Best Local Similarity 49.1%; Pred. No. 5,8e-16;
Matches 310; Conservative 0; Mismatches 298; Indels 24; Gaps 3;

```

QY 124 CAAAGATTGTGTGGTGAACCTACCAATCAGGGGGAGAGCTCCATCAGCGCCGATC 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 CAACGGTTATTTTATGTTGAATCTCAATCATGACAGCTTCTTCTATTAAGCGTG 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 184 GAGCTTACCAATCTGACGTGAGTGGCTTACCAAGCAGGAGACATCTACTTTTGGCG- 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 GATGTACCAATCATATGATGGTGGCTACCGCTGGAATATGCGCATATTTCTTCAAT 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 243 -----CGACCCACCAAGCGCGGGAAGGCAATCTCTTCAACCGGACCCAGCA 291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 CCTGACATTCAGGAATGAGATGAGCAATCATCTATCTTTTCACTGATTTCAAAATGCA 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 292 TCTCTCTCCATTTCAACCGGAATCTACAGATCTGAGAGATAGCGCGTAT---AGG 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 TATACATTCGCTTTTGTGTGATTTATGATAGCTTGAACACTTCTGTGATTTCTGACA 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 349 GACCAATCCCTCTGGGTATAGAGAACTCATTCATCCGCTCGGCGCTTGTTA---- 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 GAAATATGAGTTGGGAAATGCTCCAGTAGAGAGGCTATCTCAGCGCTTATTTATTTAC 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 -----TCAGGCGGCGACACCGCGCCCAAGCTCGTTCCCTTAATCTCATTCAGATG 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 598 AGTACGTGGTGGCACTGAGCTTCACACTGCGCTTCCCTTAATATTTGCAATCCAAATG 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 ATCTCCGAGGCGCGAGATTCATTCCTTTTGGAGGCGTCCCAATATACATTACAGC 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 658 ATTTCAGAGCAGACAGATTTCCATATTTAGAGGGAATATGCGCAGAAATTTAGTAC 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 GGGGAGTATTTCTTCCGACATGATGCTGAGCTGAGAGCTGAGGAGCTGAGGCGCAACA 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 718 AACCGAGATCTCAGCAGATCTCTAGTAAATTTACACTTGAGAAATTTGGGGAGACATT 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 TCCAGCAGTCCAGAGATCTAGGATGGCGTTTAAATACCATTTGGTGGGTATA 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 778 TCCAGTGCATTTCAAGATCTTAACCAAGAGCTTTGCTGATTCATTAACCTGCAAAA 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 TCCAGCAGTCTGAGCTGAGCAATGTCGAGAGCTGAGGAGCTGAGGAGCTGAGGAG 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838 CGTAATGTTCCAAATTCAGTGTGATGATGATATTAATCCCTATCATAGCTCTC 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 ATGTTGTTGATGATGAGGAGCAGCATCTTC 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898 ATGTTGATAGATGCGCAGCTCCACCATCTC 929
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

```

US-08-425-336-11
/ Sequence 11, Application US/08425336
/ Patent No. 5621083
/ GENERAL INFORMATION:
/ APPLICANT: Better, Marc D.
/ APPLICANT: Carroll, Stephen F.
/ APPLICANT: Studnika, Gary M.
/ TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
/ TITLE OF INVENTION: Proteins
/ NUMBER OF SEQUENCES: 140
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 6300 Sears Tower, 233 South Wacker Drive
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60606-6402
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/425,336

```

```

/ FILING DATE: 18-APR-1995
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/064,691
/ FILING DATE: 12-MAY-1993
/ APPLICATION NUMBER: US 07/901,707
/ FILING DATE: 19-JUN-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/787,567
/ FILING DATE: 04-NOV-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meyers, Thomas C.
/ REGISTRATION NUMBER: P-36,989
/ REFERENCE/DOCKET NUMBER: 31394
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312/474-6300
/ TELEFAX: 312/474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 813 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ US-08-425-336-11

```

Query Match 7.9%; Score 66.4; DB 1; Length 813;

Best Local Similarity 48.2%; Pred. No. 1.7e-12;
Matches 292; Conservative 0; Mismatches 296; Indels 18; Gaps 3;

```

QY 128 GATTGTGCTGTGTAACCAATCAGAGGGGAGAGCTGATCAGCGCCGATCGAGC 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 GTTCTGTTGTGGTACCGCTTCAATGACATGACAGTGGCGGAATAGCTATGATG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 188 TTACCAATCTGAGTGGTGGCTTACCAAGCAGGCGACCAATCTTACTTTTGGCGAGC 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 TTACAAGTTTATGTTGTTGGCTATCAAGTAAAGAAACAGATCTTACTTTTAAAGATG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 248 CACGAGCGGGCGGAAAGGATCTCTTACCGCGCACACAGATCTCTCCCATTTCA 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 CTCGAGATGCTGCTTACCAAGCGCTCTTCAAAAACACAATTTAAACAAGACTTATTTG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 308 CCGAAGCTTACACAGATCTGAGCGATAGCGCGGTCTATAGGACAGATCCCTCTGGGTA 367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 GCGGAGGATATCCCTCCGCTGGAAGTGAAGGATGATAGAGACAACAGACTTGGGCA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 368 TAGAGACATCTATCAATCCCTCTCGCGCTTCTGTTATCCAGGGCGAGCA----- 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 TTGAACCATTAAGGATTTGGCATCAAGAAACTTGATGAAATGCGATAGACAAATTATAAC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 419 CCGGCGCCCAAGCTGTTCCCTTAATTCCTCATTCAGATGATCTCGAGGCGCGAGAT 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 CACGAGATATGATAGTATGTTCTATATTTGTTATTCAAATGATGTCGAAGCAGCTGAT 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 479 TCAATCCATCTTTTGGAGGCTCGCCAAATACATTAACAGCGGGAGTCAATTTCTCCG 538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 T-----CACCTTATTTAGAAACCAATTAAGAAATCTTCAACAGAGATTCGCCGG 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 539 ACATGTATATCTGAGTGTGAGACTAGTTGGGCGCAACAATCCAGCAGATCCAGCAT 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 563 CGAATTAATCAATAGCTTGAAGATTAATGGGGTAAATCTGCTTCAAGATCCGAGCAT 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 599 C---TACGATGGGCTTTTAAATTAACCATTTGCTGGGTATATCCACCGTAACTTCG 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 623 CAGGTGCAATGGAATGTTTTCGAGAGGAGTGAATTTGGAACGTGCAAAATGGCAAAAT 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 656 TGACGTTGAGCAATGTTGCGAGCGTATCGCAGCTTAGCATCATGTTGTTGTATGTA 715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 ACTATGTACGCGAGTGTATCAATTAACCCAAATAGACACCTTGAAGTTGCTGATA 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 716 GGGACC 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Thu Mar 27 07:06:57 2003

us-09-627-165e-15.rni

Page 11

Db 743 MAGATC 748

Search completed: March 26, 2003, 19:43:05
Job time : 58.4724 secs

•
•
•
•

	Query Match	77.0%	Score 644.2;	DB 10;	Length 756;
	Best Local Similarity	92.7%;	Pred. NO. 1e-217;	Mismatches 53;	Indels 0; Gaps
	Matches 676;	Conservative	0;		
OY	1 CACGAGCAGCGGGGACGAATATTTCGGGTCTACAGCTTCCGCAATTATGTCTCA	60			
Db	28 CACCAGCACACGGTGAAAGAAATTTCGGTTCTACAGCTTCCGCAATTATGTCTCA	87			
OY	61 AGCGAAGCTTTTCCAATGAGATAACAACCTGTGGGTAGTCTGCAGTCCCCTTCGGAT	120			
Db	88 AGCGAAGCTTTTCCAATGAGATAACAACCTGTGGGTAGTCTGCAGTCCCCTTCGGAT	147			
OY	121 GCGCCAAGAATTTGTGTTGGTGAACATCCCAATAGGGGGAGACTCGATCACGGCCGCC	180			
Db	148 GCCCAAAGATTTGTCTTGTTGGAGCTCACCAACAGGGGGAGACTCGATCACGGCCGCC	207			

OY	181	ATCGACGTTACCAATCTTACGTGTGGCTTACCAAGACGGCCCAATCCTACTTTTG	2140
Db	208	ATCGACGTTACCAATCTTACGTGTGGCTTACCAAGAGGGCCAACTCTACTTTTG	267
OY	241	CGCGACGACCAAGACGGCGGGAAAGGATCTCTTCAACGGCCACACCAAGATCCTCTC	3007
Db	268	CGCGACGACCAAGCGGGCGGGAAAGGATCTCTTCAACGGCCACACCCGATCCTCTC	3277
OY	301	CCATTACACCGGAGCTACACAGATCTGGAGCGATACGCCGGCTATYAGGACCAAGATCCT	3606
Db	328	CCATTCAACGGAGCTACCCCTGATCTGGAGCGATACGCCGGCATYAGGACCAAGATCCT	3877
OY	361	CTGGGTATAGAGGAACATTCATTCGCTCGTGGCGGCTTGTTATCCAGCGCGAGACAC	4207
Db	388	CTCGGTATAGACCAATCTATTCATTCGTCACGGCGCTGTGTTTCCGGCGCGACAGACG	4477
OY	421	CGGGCCCAAGCTCGTCTCCCTTATATCCTCATTCAGATGATCTCCAGAGCGCGAGATTC	4807
Db	448	CGTACCCAAAGCTCGTGTGATTTTATCCTCATTCAGATGATCTCCAGAGCGCGCAATTC	5077
OY	481	AATCCCATCTTTTGGAGGGCTCGCCAAATACATTAAACGCGGGGAGTCATTTCTCCGAC	5407
Db	508	AATCCCATCTTATGGAGGGCTCGCCCAATACATTAAAGTGGGGGCTCATTTCTGCGAGAC	5677
OY	541	ATGTACATGCTCGAGGCTGGAGACATAGTGGGGCCCAAAATTCACGCAAGTCCAGACTCT	6007
Db	568	GTTACATGCTGGAGGCTGGAGACAGATGGGGCCCAAAATTCACGCAAGTCCAGACTTCA	6277
OY	601	ACGGATGGCGTTTTTAATTAACCCATTCGCTGGGTATATCCACCGGTACTCTCGTAGCG	6607
Db	628	ACCGATGGCGTTTTTAATTAACCCATTCGCTGGGTATATCCCGCGGTAACTTCGTAGCG	6877
OY	661	TTTGACCAATGTTCCGCGACGTGATCGCCAGCTTAAGCCATGTTGTTGTATGTAGGAC	7207
Db	688	TTTGACCAATGTTCCGCGACGTGATCGCCAGCTTGCGCATCATGTTGTTGTATCGGAGAG	7477
OY	721	CGACCATCT 729	
Db	748	CGGCGATCT 756	

```

RESULT 2
US-09-347-064-1
; Sequence 1, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jürgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Viscum album
; US-09-347-064-1

```

Query Match	76.5%	Score 640.2;	DB 10;	Length 762;
Best Local Similarity	92.7%;	Pred. No. 2,7e-216;		
Matches 672;	Conservative	0;	Mismatches 53;	Indels 0;
			Gaps	0;
1	CATCAGACGACGGGGCAGCAATATTTCCGGTTTCATCAGCTTCCGAGATATATGTCCTCA	60		

D	b	34	CACCAACCAACCGGGAGAAATATTTCCGGTTATCAACGCTTCCGAGATTATGTCTCA	93
Q	y	61	AGCGGAAGCTTTTCCAAATGATATACACTCTTGGCGTAGTACGATATCCCGCTCGGAT	120
D	b	94	AGCGGAAGCTTTTCCAAATGATATACACTCTTGGCGTAGTACGATATCCCGCTCGGAT	155
Q	y	121	GGCGAAAGATTGTGTGTGTGGAACTACCAATCAGGGGGAGACTCGATCAGCGCGCC	180
D	b	154	GGCGAAAGATTGTGTGTGTGGAGCTACCAACAGGGGGAGACTCGATCAGCGCGCC	211
Q	y	181	ATGAGGTTACCAATCTGAGTGGTGGCTTACCAAGCAGGCAATCCACTTTTG	244
D	b	214	ATGAGGTTACCAATCTGAGTGGTGGCTTACCAAGCAGGCAATCCACTTTTG	277
Q	y	241	CGGAGCGACAGAGCGGGCGGAAAAGGCACTCTTCACGGGCAACGAGATCCTCTC	300
D	b	274	CGGAGCGACAGAGCGGGCGGAAAAGGCACTCTTCACGGGCAACGAGATCCTCTC	333
Q	y	301	CCATTACCGGAAGCTACACAGATCTGGAGCGATAGCGGCTCATAGGAGCCAGATCCCT	366
D	b	334	CCATTACCGGAAGCTACACAGATCTGGAGCGATAGCGGCAATAGGAGCCAGATCCCT	399
Q	y	361	CTGGGATAGAGAACTCATTCATCCGCTCGGGCGCTTCGTTATCCAGGGCGGACGACC	422
D	b	394	CTGGGATAGAGAACTCATTCATCCGCTCGGGCGCTTCGTTATCCAGGGCGGACGACC	455
Q	y	421	CGGGGCCAAGCTCGTCCCTTAATATCCATTCAGATGATCTCGAGGCGCGAGATTC	480
D	b	454	CGTACCCAGAGCTGTTGCATTATTAATCTCATTCAGATGATCTCGAGGCGCGAGATTC	513
Q	y	481	AATCCATCTTTTGGAGGGGCTCGCCAAATACATTAAAGCGGGGAGTCATTTTCCGAC	540
D	b	514	AATCCATCTTTTGGAGGGGCTCGCCAAATACATTAAAGCGGGGAGTCATTTTCCGAC	573
Q	y	541	ATGTAATCTGAGACTGGAAGACTATGTTGGGGCAACATTCACAGCAAGTCCAGAGCT	600
D	b	574	GTGTACATCTGAGACTGGAAGACTATGTTGGGGCAACATTCACAGCAAGTCCAGATTC	633
Q	y	601	ACGAGTGGGTTTTTAATTAACCAATTCGGTTGGGTATATCCACGGTAACCTTGTCAGC	660
D	b	634	ACGAGTGGGTTTTTAATTAACCAATTCGGTTGGGTATATCCACGGTAACCTTGTCAGC	693
Q	y	661	TTGAGCAATGTTCCGACGTGATCGCCAGCTTATAGCATCATGTTGTATGTAGGGAC	720
D	b	694	TTGAGCAATGTTCCGACGTGATCGCCAGCTTATAGCATCATGTTGTATGTAGGGAG	753
Q	y	721	CGACC 725	
D	b	754	CGCCC 758	

```

RESULT 3
US-09-785-921A-1
; Sequence 1, Application US/09785921A
; Patent No. US20020084334A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: LIT-PI-529
; CURRENT APPLICATION NUMBER: US/09/785,921A
; CURRENT FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 1
; LENGTH: 1698
; TYPE: DNA
; ORGANISM: Ricinus communis
; US-09-785-921A-1

```

Query Match	9.48;	Score 78.8;	DB 10;	Length 1698;
Best Local Similarity	49.28;	Pred. No. 2.7e-17;		
Matches 311;	Conservative 0;	Mismatches 297;	Indels 24;	Gaps 3;


```

Db      CCTGACATTCAGGAGATGAGAGACATTCATCTTTTCACGATGATTCACAAATGCA 447
QY      388
QY      232 TCCTCTCCATTCACCGGAGCTACACAGATCTGGAGCATACGCCGTCAT---AGG 348
Db      448 TATACATTCGCTTTGGGTATATATATAGACTTGAACAACCTGGCTGATATCTGAGA 507
QY      349 GACGAGATCCCTCTGGGTATAGAGAACTCAATCCGCTGGGCTTCGTTA---- 404
Db      508 GAAATATGAGATTGGGAAATGCTCCACTAGAGAGCTATCTCGCCCTTATATATAC 567
QY      405 -----TCCAGGGCGGAGACCCCGGCCCAAGCTCGTTCCCTTAATATCCATTCAGATG 459
Db      568 AGTACTGGGCACTACAGCTTCCAACTCTGGCTGCTCTTATATATTTGCATCCAAATG 627
QY      460 ATCTCCGAGGCGCGAGATTCATCCATCTTTTGGAGGGCTCGCCCAATPATATTAACGC 519
Db      628 ATTTGAGAGAGAGCAAAATTCCTCAATATTTAGGGAGAAATGCGCAGAGATTAAGTAC 687
QY      520 GGGAGTCAATTTCTCCGACATGTACATGCTCGAGCTGAGAGACTAGTTGGGCCAACAA 579
Db      688 AACCGAGATCTGCACACAGATCTAGCGTATATACACTTGAATATAGTTGGGGAGACTT 747
QY      580 TCCACGCAAGTCCAGAGCTACAGGATGCGCTTTTAAATAACCAATTCGGTTGGGTATA 639
Db      748 TCACATGCAATTCAGAGATCTACCAAGAGCCTTTGCTAGTCCAAATTCACATGCAAGA 807
QY      640 TCCACCGGTAATCTGCTGACCTTGAGCAATGTCCGACAGTGATGCGCAGCTTACGATC 699
Db      808 CGTAAATGTTCCAAATTCAGTGTAGTGTAGATATATATATATCCATCATAGCTCTC 867
QY      700 ATGTTGTTGTATGTAGGAGACCATCTTC 731
Db      868 ATGGTATATAGATGGGCACTCCACATCTGC 899

```

RESULT 6 US-09-785-921A-10

```

; Sequence 10, Application US/09785921A
; Patent No. US20020094334A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; FILE REFERENCE: LIT-PI-529
; CURRENT APPLICATION NUMBER: US/09/785,921A
; CURRENT FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 10
; LENGTH: 1805
; TYPE: DNA
; ORGANISM: R1c1nus communis
US-09-785-921A-10

```

Query Match 9.0%, Score 75.6; DB 10; Length 1805;
Best Local Similarity 48.9%; Pred. No. 3.7e-16;
Matches 309; Conservative 0; Mismatches 299; Indels 24; Gaps 3;

```

QY      124 CAAGATTTGTTGTTGGTGAATCTCACCAATCAGGGGAGAGACTGCATACGCCGCCATC 183
Db      276 CAACGGTTATTTTATGTTAGTCACTCAATATCATGAGAGCTTTGTTACATTAAGCGCTG 335
QY      184 GACGTTACCAATCTGTACGTGCTGCTTACCAAGAGCGGAGCAATCTTACTTTTGGG- 242
Db      336 GATGTACCAATGATATGTGCTGCTACCGCTGTGAAATAGGCAATATTTCTTTTAT 395
QY      243 -----CGACGACACAGAGCGCGCGGAAAGCATCTTTACCGGAGACCAACAGA 291
Db      396 CCGTACCAATCAGAGATGAGAGCAAGCAATCACTCTTTTACATGATGTTCAAAATGCA 455
QY      292 TCCTCTCTCCATTCACCGGAGATGACAGATCTGGAGCGATACGCCGCTCAT---AGG 348

```

```

Db      456 TATACATTCGCTTTGGGTATATATATAGACTTGAACAACCTGCTGTAATCTGAGA 515
QY      349 GACGAGATCCCTCTGGGTATAGAGAACTTCATTCATCCGCTCGCGCTGTGTTA---- 404
Db      516 GAAATATGAGTTGGGAAATGCTCCACTAGAGAGCTATCTCACGCGCTTATATTAATGAC 575
QY      405 -----TCCAGGGCGGAGACCCCGGCCCAAGCTCGTTCCCTTAATATCCCTATTCAGATG 459
Db      576 AGTACTGGTGGCACTACAGCTTCCAACTCTGGCTGCTCTTATATATTTGCATCCAAATG 635
QY      460 ATCTCCGAGGCGCGAGATTCATCCATCTTTTGGAGGGCTCGCCCAATPATATTAACGC 519
Db      636 ATTTGAGAGAGAGCAAAATTCCTCAATATTTAGGGAGAAATGCGCAGAGATTAAGTAC 695
QY      520 GGGAGTCAATTTCTCCGACATGTACATGCTCGAGCTGAGAGACTAGTTGGGCCAACAA 579
Db      696 AACCGAGATCTGCACACAGATCTAGCGTATATACACTTGAATATAGTTGGGGAGACTT 755
QY      580 TCCACGCAAGTCCAGAGCTACAGGATGCGCTTTTAAATAACCAATTCGGTTGGGTATA 639
Db      756 TCACATGCAATTCAGAGTCTACCAAGAGCCTTTGCTAGTCCAAATTCACATGCAAGA 815
QY      640 TCCACCGTAACTTCTGACCTTGAGCAATGTCCGACGCTATGCGCAGCTTACGATC 699
Db      816 CGTAAATGTTCCAAATTCAGTGTAGTGTAGATATATATATCCATCATAGCTCTC 875
QY      700 ATGTTGTTGTATGTAGGAGACCATCTTC 731
Db      876 ATGGTATATAGATGGGCACTCCACATCTGC 907

```

RESULT 7 US-09-765-527-246

```

; Sequence 246, Application US/09765527
; Patent No. US20020006638A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; TITLE OF INVENTION: Methods for Recombinant Microbial Production of
; Fusion Proteins and BPI-Derived Peptides
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,527
; FILING DATE: 18-Jan-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/621,803
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27129/33199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 246:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 813 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```



```

;          FEATURE:
;
;          NAME/KEY: misc.feature
;
;          OTHER INFORMATION: "gelonin"
;
;          SEQUENCE DESCRIPTION: SEQ ID NO: 246
US-09-765-527-246

```

Query Match	7.9%;	Score 66.4;	DB 10;	Length 813;
Best Local Similarity	48.2%;	Pred. No. 4.6e-13;		
Matches 292; Conservative	0;	Mismatches 296;	Indels 18;	Gaps 3;

QY	128	GATTTGGTGGTGGAACTCAACCAATAGGGGGAGACTGATACGGCCCATCGACG	187
Db	149	GTTCGTGTTGGTAGCCCTTTCAAATGACAAATGACAGTTGGGGAATAGCTATAGATG	208
QY	188	TTACCAATCTGTACGTGGTGGCTTACCAAGACAGCGCAATCCCTACTTTTGGCGACG	247
Db	209	TTACAAAGTGTTAATGTGTGGGCTATCAAGTAGAAGAACAACTCTTACTTTTAAAGATG	268
QY	248	CACCAAGCGGGCGGAAAGCAATCTTTCACCGGACACCAACAGATCCTCTCCCAATTCA	307
Db	269	CTCCAGATGCTGCTTACGAAGGCCCTCTTCAAAAACAACTTAAACAGACTTCATTG	328
QY	308	CCGGAACCTACAGATCTGGAGGATAGCCCGGTCAATAGGACAGATCCCTGGGTA	367
Db	329	GGGGACACTTCCCTCGCTGGAAAGGTGAGAAGGCATATAGAGACAAACGACTTGGCA	388
QY	368	TAGAGAACTCATTCATCCGCTCGCGCTTCGTGTTATCCAGCGCGAGCA-----	418
Db	389	TTGAACCATTAAGGATGGCATTAAGAACTTATGAAATTCGGATAGACAATTAATTAAC	448
QY	419	CCGGGGCCCAAGTCTGTTCCCTTAATAATCTCATTCAGATGATCTCGAGGCCGAGAT	478
Db	449	CACGGGAGATGATGATGTTCTCTATATGGTTGTTATTCAAATGAGTCTGGAAGCAGCTGAT	508
QY	479	TCAATCCCATCTTTGGAGGGCTCGCCAAATACATTAACAGGGGGAGTCAATTCCTCCG	538
Db	509	T-----CACCTTATTTGGAACCAAAATTAGAAATTAACCTTCAACAGAGATGTGCCGG	562
QY	539	ACATGTACATGCTCGAGCTGGAGACTAGTTGGGGCCCAACATTCACGCAAGTCCACAGT	598
Db	563	CGAATATATCAACCTCTTGAGAAATAAATGGGGTAAACTCTGCTTCAGATCCGACAT	622
QY	599	C---TAAAGATGGCGTTTAAATAACCAATTCGTTGGGTATATCCACCGGTAACCTGG	655
Db	623	CAGGTGAATGGAATGTTTTCGAGGCGAGTTGAAATTTGGAACGTGCAGAAATGGCAAAAT	682
QY	656	TGACGTTGAGCAATCTTCGGAGCGTATGCGCCAGCTTAGAGATCATGTTGTGATGTA	715
Db	683	ACTATGCAACCGAGTGTATCAAGTAATAAACCCAAATATGCACTCTTGAAGTGTGTGATA	742
QY	716	GGGACC 721	
Db	743	AAGATC 748	

```

1      RESULT 8
2      US-09-765-527-258
3      ; Sequence 258, Application US/09765527
4      ; Patent No. US20020006638A1
5      ;
6      GENERAL INFORMATION:
7      ;
8      APPLICANT: Better, Marc D.
9      ;
10     TITLE OF INVENTION: Methods for Recombinant Microbial Production of
11     ;
12     ; Fusion Proteins and BPI-Derived Peptides
13     ;
14     NUMBER OF SEQUENCES: 265
15     ;
16     CORRESPONDENCE ADDRESS:
17     ;
18     ADDRESSSEE: Marshall, O'Toole, Gerstein, Murray & Borun
19     ;
20     STREET: 6300 Sears Tower, 233 South Wacker Drive
21     ;
22     CITY: Chicago
23     ;
24     STATE: Illinois
25     ;
26     COUNTRY: United States of America
27     ;
28     ZIP: 60606-6402
29     ;
30     COMPUTER READABLE FORM:
31     ;
32     MEDIUM TYPE: Floppy disk

```

```

1      COMPUTER: IBM PC compatible
2      OPERATING SYSTEM: PC-DOS/MS-DOS
3      SOFTWARE: Patentin Release #1.0, Version #1.25
4
5      CURRENT APPLICATION DATA:
6      APPLICATION NUMBER: US/09/765,527
7      FILING DATE: 18-Jan-2001
8
9      PRIOR APPLICATION DATA:
10     APPLICATION NUMBER: 08/621,803
11     FILING DATE: <Unknown>
12
13     ATTORNEY/AGENT INFORMATION:
14     NAME: Borun, Michael F.
15     REGISTRATION NUMBER: 25,447
16     REFERENCE/DOCKET NUMBER: 27129/33199
17
18     TELECOMMUNICATION INFORMATION:
19     TELEPHONE: 312/474-6300
20     TELEFAX: 312/474-0448
21     TELEX: 25-3856
22
23     INFORMATION FOR SEQ ID NO: 25:
24
25     SEQUENCE CHARACTERISTICS:
26     LENGTH: 955 base pairs
27     TYPE: nucleic acid
28     STRANDEDNESS: single
29     TOPOLOGY: linear
30
31     MOLECULE TYPE: protein
32
33     FEATURE:
34     NAME/KEY: CDS
35     LOCATION: 66..944
36
37     FEATURE:
38     NAME/KEY: misc_feature
39     LOCATION: residues 1-65
40     OTHER INFORMATION: /label= EcoRI
41     /note="residues 1-65 comprise EcoRI site to beginning of pel
42     B."
43
44     FEATURE:
45     NAME/KEY: misc_feature
46     LOCATION: AA 1-22
47     OTHER INFORMATION: /label= pel B
48     /note="pel B is the leader sequence from the pectate lyase
49     gene of Ewinlia caratovora."
50
51     FEATURE:
52     NAME/KEY: misc_feature
53     LOCATION: AA 23-273
54     OTHER INFORMATION: /label= "gelonin"
55     /note="gelonin (see U.S. Patent No. US20020006638A1 5,416,202)."
56
57     FEATURE:
58     NAME/KEY: misc_feature
59     LOCATION: AA 274-275
60     OTHER INFORMATION: /label= BagI
61     /note="BagI cloning site."
62
63     FEATURE:
64     NAME/KEY: misc_feature
65     LOCATION: AA 276-279
66     OTHER INFORMATION: /label= cleavage linker
67     /note="Ala-Leu-Asp-Pro linking sequence with Asp-Pro cleavage
68     site."
69
70     FEATURE:
71     NAME/KEY: misc_feature
72     LOCATION: AA 280-293
73     OTHER INFORMATION: /label= peptide sequence
74     /note="BPI-derived peptide."
75
76     FEATURE:
77     NAME/KEY: misc_feature
78     LOCATION: residues 945-954
79     OTHER INFORMATION: /label= XhoI
80     /note="residues 945-955 comprise stop codon and XhoI site."
81
82     SEQUENCE DESCRIPTION: SEQ ID NO: 25:
83     1S-09-765-527-258

```

0y 128 GATTTCGCTGGGAGACCTACCAATCAGGGGGAGACTCGATACGAGCCGCATCGAGC 187

Query Match	7.88;	Score 65;	DB 10;	Length 955;
Best Local Similarity	47.8%;	Pred. No. 1.6e-12;		
Matches 296;	Conservative 0;	Mismatches 305;	Indels 18;	Gaps 3;

OY	308	CCGGAAGCTACACAGATNTGGAGCGATACGCCGCTCATAGGAGCAGAGATCCCTCTGGGTA	367
Db	460	GGGAGCGATATCCCTCGTGGAGAGTGGAAGGCATATAGAGACAAACAGACTTGGGCA	519
OY	368	TAGAGAACTCATTCATCCCTCTCGGGCGCTCGTTATCCAGGGGGGAGCA-----	418
Db	520	TTGAAACCATTTAAGGATTGGCATCAAGAACTTGATGAAAAATGCCGATAGACAAATTTATTAAC	579
OY	419	CCCGGGCCCAAGCTGCTTCCCTTAAATCCTCATAGATGATCCGAGGCCGCGAGAT	478
Db	580	CAACGAGATAGCTAGTACTCTATTGGTTGTTATTCAAAATGGTGTCGAAGCAGCTCGAT	639
OY	479	TCAATCCCATCTTTTGGAGGGCGTCCCAATATCATTAACACCGGGGAGTCACTTTCTCCCG	538
Db	640	T-----CACTTTATTGAGAACCAAAATTAAGAAATTACTTCAACAGAGATATCGGCCGG	693
OY	539	ACATGTACATGCTCGAGCTGGAGACTAGTTGGGGCCCAACATCCACGACAGTCCAGCAGT	598
Db	694	CGAATTAATACATCAGGCTTGAGAAATTAATGGGGTAACTCTGCTCCAGATCCGACAT	753
OY	599	C---TAAGGATGGCGTTTATATATACCCATTTGGTGGGTATATCCACGCGATACCTTG	655
Db	754	CAGGTGCAAAATGGAATTTTTCGGAGGCACTTAATTTGAAACGTGCAAAATGGCAAAAAT	813
OY	656	TGACGTTGAGCAATGTTGCGGAGCTGATTCGACAGTATGAGATCATGTTTGTATGTA	715
Db	814	ACTATGTACGCGCAGTTGATCATCAAGTAATAACCCAAAATAGCACTCTTGAAGTTGTCGATA	873
OY	716	GGGACGACCATCTTCTCTC	734
Db	874	AAGATCTTAATCGGCCG	892

US-09-765-527-250
; Sequence 250, Application US/09765527
; Patent No. US2002000638A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 250:
SEQUENCE CHARACTERISTICS:
LENGTH: 1072 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

	?	TOPOLOGY:	Linear
	?	MOLECULE TYPE:	protein
	?	FEATURE:	
	?	NAME/KEY:	CDS
	?	LOCATION:	66..1061
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	residues 1-65
	?	OTHER INFORMATION:	/label= EcoRI
	?	/note="Residues 1-65 comprise EcoRI site to beginning of pel B."	
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	AA 1-22
	?	OTHER INFORMATION:	/label= pel B
	?	/note="pel B is the leader sequence from the pectate lyase gene of Erwinia carlotovora."	
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	AA 23-273
	?	OTHER INFORMATION:	/label= "gelonin"
	?	/note="gelonin (see U.S. Patent No. US2002000638A1 5,416,202)."	
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	AA 274-276
	?	OTHER INFORMATION:	/label= EagI
	?	/note="EagI cloning site."	
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	AA 277-296
	?	OTHER INFORMATION:	/label= SLT linker
	?	/note="SLT from shiga-like-toxin gene."	
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	AA 297-298
	?	OTHER INFORMATION:	/label= FspI/ScaI
	?	/note="FspI and ScaI cloning sites."	
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	AA 299-302
	?	OTHER INFORMATION:	/label= cleavage linker
	?	/note="Ala-Leu-Asp-Pro linking sequence with Asp-Pro cleavage site."	
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	AA 303-332
	?	OTHER INFORMATION:	/label= peptide sequence
	?	/note="BPI-derived peptide."	
	?	FEATURE:	
	?	NAME/KEY:	misc_feature
	?	LOCATION:	residues 1062-1072
	?	OTHER INFORMATION:	/label= XhoI
	?	/note="Residues 1062-1072 comprise stop codon and XhoI site."	
	?	SEQUENCE DESCRIPTION:	SEQ ID NO: 250;
	?	US-09-765-527-250	
	?	Query Match	7.8%; Score 65; DB 10; Length 1072;
	?	Best Local Similarity	47.8%; Pred No. 1,6e-12;
	?	Matches 296: Conservative	0; Mismatches 305; Indels 18; Gaps 3;
OY	128	GATTGTGTTGGTGGAACCTCACCAATCAGGGGGAGACTCGATTCACGGCCGCATCGACG	187
Db	280	GTTTTGTTGGTTGTAACCGCTTCATAGCAATGACAGACAGTTGGCGAATAAGCTATAGATG	339
OY	188	TTACCACATCTGAGCTGGGGCTTCCCAAGCAGCGGACCAATCTTACTTTTGGCGGACG	247
Db	340	TTACAGATGTTTATGTGTGGGTCTTCACAAGTAGAAGAAGATCTTACTTTTAAAGATG	399
OY	248	CACGACGCGCGGAAAGCATCTCTTCACGCGGACACAGATCTCTGCCATCA	307
Db	400	CTCCAGATGCTGCTTCAGCAAGGCGCTTCCAAMACAAATTAAMCAAGACTTCATTG	459
OY	308	CCGGAAGCTACACAGATCTGGAAGCATACGCCGGTCAATAGGAGACCAGATCCCTCTGGGTA	367

[illegible]

```

RESULT 11
US-09-464-099A-2
; Sequence 2, Application US/09464099A
; Patent No. US20020168680A1
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: GLYCOPHOSATE TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASES
; FILE REFERENCE: 11899, 0175, CWO501 MOBT:175-2
; CURRENT APPLICATION NUMBER: US/09/464, 099A
; CURRENT FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: US 09/137, 440
; PRIOR FILING DATE: 1998-08-20
; PRIOR APPLICATION NUMBER: US 08/833, 485
; PRIOR FILING DATE: 1997-04-07
; PRIOR APPLICATION NUMBER: US 08/306, 063
; PRIOR FILING DATE: 1994-09-13
; PRIOR APPLICATION NUMBER: US 07/749, 611
; PRIOR FILING DATE: 1991-08-28
; PRIOR APPLICATION NUMBER: US 07/576, 537
; PRIOR FILING DATE: 1990-08-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1982
; TYPE: DNA
; ORGANISM: Agrobacterium sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (62)..(1426)
US-09-464-099A-2

Query Match 4.6%; Score 38.8; DB 9; Length 1982;
Best Local Similarity 52.5%; Pred. No. 0.0042;
Matches 85; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 137 TGTGCACTCACCATCAGGGGGAGACTGCATCAGCGCCGCAATGAGCTTACCAATC 196
1101 TGGAGAGATCCGCGTCGCAAGGAAGCGACCGCCCTTCGGCGCTGCCAATGGCCCTCAGC 1160

```

OY	197	TCGACGTGCTGCTTACCAACGACGGGACCAACCTACTTTTGGGCGACGACACG	256
Db	1161	TCGATGGGTGATGGATGAGAGGGGACGACGTGCTGCTGTGGCGGCGGCGCTGAC	1220
OY	257	GGCGGAGAAAGGCATCTTCTACCGGCGCACCGATGCTCTC	298
Db	1221	GCAAGGGGCTGCGCAACGCGCTCGGGGGCGCGCTGCGCACCC	1262

```

RESULT 12
US-09-861-696-2
/ Sequence 2, Application US/09861696
/ Patent No. US2002007053A1
/
/ GENERAL INFORMATION:
/ APPLICANT: Barry, Gerard F.
/ APPLICANT: Kishore, Ganesh M.
/ APPLICANT: Padgett, Stephen R.
/ APPLICANT: Stallings, William C.
/ TITLE OF INVENTION: GLYPHOSATE TOLERANT 5-ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHAS
/ FILE REFERENCE: 11899-0175-CUS04 MOBT-175-4
/ CURRENT APPLICATION NUMBER: US/09/861,696
/ CURRENT FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 09/137,440
/ PRIOR FILING DATE: 1998-08-20
/ PRIOR APPLICATION NUMBER: US 08/833,485
/ PRIOR FILING DATE: 1997-04-07
/ PRIOR APPLICATION NUMBER: US 08/306,063
/ PRIOR FILING DATE: 1994-09-13
/ PRIOR APPLICATION NUMBER: US 07/749,611
/ PRIOR FILING DATE: 1991-08-28
/ PRIOR APPLICATION NUMBER: US 07/576,537
/ PRIOR FILING DATE: 1990-08-31
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: Patencin version 3.0
/
/ SEQ ID NO 2
/
/ LENGTH: 1982
/ TYPE: DNA
/
/ ORGANISM: Agrobacterium sp.
/
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (62)..(1426)
/
US-09-861-696-2

```

```

Query Match          4.6%; Score 38.8; DB 10; Length 1982;
Best Local Similarity 52.5%; Pred. No. 0.0042;
Matches      85; Conservative      0; Mismatches 77; Indels      0; Gaps

```

QY	137	TGTTGGAATCTACACCAATAGGGGGAGAGACTGTGATCAGCGGGCGGCAATGAGCTTACCAATC	196
DB	1101	TGAGAAACTCCGCGCTCAAGGAAGGACCGCCCTCTCGGCGGTGGCCATGCTTAAACG	1160
QY	197	TGTACTGTGTGCTTACCAAGAGCGGCGCAATCTCTATTTTTGGCGAGCAGCACAGAG	256
DB	1161	TCATATGCGCTGATTGTGCATGTGGGCGGAGCGTGCCTGTCTGTCGGCGGCGCCCTTGACG	1220
QY	257	GGCGGGAAGGCAATCTTTACACCGGCGACACCAAGATCTCTTC	298
DB	1221	GCAAGGGGCTCGGCAACGCTTGGGCGCGCCGTGCGCACCC	1262

```

RESULT 13
US-09-347-064-11
; Sequence 11, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe viscum
; FILE REFERENCE: 09282-5

```

```

; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Viscum album
US-09-347-064-11
```

```

Query Match 4.5%; Score 38; DB 10; Length 48;
Best Local Similarity 89.1%; Pred. No. 0.0014;
Matches 41; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```

QY 730 TCCTCCGAGCTGCGCTATTGGCCGCTGTCATACGACCCGCTTGG 775
Db 1 TCCTCTGAGGTGCGCTATTGGCCGCTGTCATACGACCCGCTGATAG 46
```

```

RESULT 14
; Sequence 584, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See file wrapper or palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 584
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-584
```

```

Query Match 4.3%; Score 35.6; DB 9; Length 708;
Best Local Similarity 12.3%; Pred. No. 0.035;
Matches 48; Conservative 108; Mismatches 233; Indels 0; Gaps 0;
```

```

QY 125 AAGATTGTGTGGTGAACCAATCAGGGGAGAGAGCTGATCAGGCGGCATG 184
Db 663 A.RCTTYR.YG..BKC.D.N.HR.A.AC..C.S.....CCC.T.C.S.SBSBCTYMR 604
QY 185 ACCTTACATCTGTACGTGTGGCTTACCAAGCAGGCAACCAATCTTTTGGCGG 244
Db 603 RCNA.HNA.AC...B.HSDHADB..CYH.YSCA....CYSSCHSSC..YK.NM..S 544
QY 245 ACGCAGCAGGCGGCGGAAAGGATCTTTCACCGGACGACCAAGATCTTCCCAT 304
Db 543 .STSTMSDYWG..TBCA.BK..AAATTTCTB.YBA..HYTC.ST.TCC.C.YNYTSGN 484
QY 305 TCACCGGAACTACACAGATCTGGAGCGATACCGCGCTATGAGGACGACATCCCTGTG 364
Db 483 .STS.HS..ACS.D.BS..C.C.CCYS..T..A.B..YC.SCT...C.HMSBPTH.T.S 424
QY 365 GATAGAGAACTCATTCATCCGTCGTCTGGCGCTTCGTTATCCAGGCGGACACCGGG 424
```

```

Db 423 T.B.G..YT.CY..T.TSD.TABB..TNBA.YT.TTCY.CS.CT.TYBRRQTH..CBS 364
QY 425 CCAAGCTCGTTCCTTATATATCCATGATGATTCGAGCGCGGAGATTCATC 484
Db 363 ..Y.ATCSC.B.TCCYAMABY...NCYWTGS.STH.CHBBSR..MT....C.BH.TA 304
QY 485 CCATCTTTGGAGGCGCGCCCATACATT 513
Db 303 SYNTSCTNSB.GSB..A..STHTTB..Y 275
```

```

RESULT 15
US-09-864-761-13457/C
; Sequence 13457, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; FILE REFERENCE: Aecmice-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13457
; LENGTH: 526
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011449.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
```

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
US-09-864-761-13457

Query Match 4.1%; Score 34.4; DB 10; Length 526;
Best Local Similarity 59.0%; Pred. No. 0.081;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 528 ATTTCCTCCGACATGCTGAGCTGAGACTAGTTGGGGCCACATCCACGCA 587
||| ||||| | |||| | | | | | | | | | | | | | | |
Db 382 ATTACTCCCATTAATACATCATTAAGTCGAAATTCCTAGTGAACCATCCATCGTA 323
OY 588 AGTCCAGCAGCTACGATGGCGTTTAAATACCATT 627
||| | | | ||| | ||||| ||||
Db 322 AGTCAGGGAATGCTGATTAAGCTTTAATAGATCATTT 283

Search completed: March 26, 2003, 19:45:49
Job time : 103.636 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 16:07:41 ; Search time 1425.84 Seconds
(without alignments)
9507.085 Million cell updates/sec

Title: US-09-627-165E-15

Perfect score: 837
Sequence: 1 catcagacgacgagcagca.....aacccaccgctgcgacgtcga 837

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_iny:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48.8	5.8	405	14	T24255
2	43.6	5.2	479	12	BG789629
3	43.6	5.2	558	10	BE606625
4	42.8	5.1	811	10	BE034055
5	42.2	5.0	837	10	AM053634
6	42.2	5.0	1021	10	BE036639

7	41.6	5.0	672	10	AM053585
8	40.4	4.8	779	14	BM816277
9	40.4	4.8	913	14	BM816276
10	39.8	4.8	586	14	BQ765173
11	38.8	4.6	463	10	BE425390
12	38.8	4.6	589	9	AJ436193
13	38.8	4.6	621	10	AV836389
14	38.8	4.6	623	14	BQ767985
15	38.8	4.6	626	10	BE060235
16	38.8	4.6	631	13	BQ767985
17	38.8	4.6	641	14	BQ739975
18	38.8	4.6	700	9	AL505987
19	38.8	4.6	867	13	BI953600
20	38.4	4.6	420	9	AJ466017
21	38.4	4.6	540	12	BG367680
22	38.2	4.5	296	12	BG043201
23	37.4	4.5	829	17	BH526892
24	37.2	4.4	408	10	AM047216
25	37.2	4.4	561	10	BE426610
26	37.2	4.4	565	10	BE470946
27	36.4	4.3	557	14	BQ764896
28	36.4	4.3	610	14	BQ764482
29	36.4	4.3	788	12	BG310389
30	35.6	4.3	443	13	BM110994
31	35.6	4.3	545	10	BE511221
32	35.6	4.3	686	9	AI890522
33	35.6	4.3	993	10	BE035039
34	35.4	4.2	699	13	BI684857
35	35.4	4.2	385	13	BM136974
36	35	4.2	1094	11	AY105813
37	34.8	4.2	453	17	CNS03975
38	34.8	4.2	611	14	BQ490182
39	34.6	4.1	565	13	BI396105
40	34.6	4.1	620	9	AJ451296
41	34.6	4.1	661	9	AJ452361
42	34.6	4.1	678	9	AJ450811
43	34.4	4.1	368	14	BQ171562
44	34.2	4.1	888	12	BG369039
45	34	4.1	186	10	BA416990

ALIGNMENTS

RESULT 1
T24255
LOCUS 405 bp mRNA linear EST 28-JUL-1995
DEFINITION csl345 lambdazAPST Rictinus communis CDNA clone pcrl345, mRNA
ACCESSION T24255
VERSION T24255.1 GI:689074
KEYWORDS EST.
SOURCE castor bean.
ORGANISM Rictinus communis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Rictinus.
1 (bases 1 to 405)
vandeloo,F.J., Turner,S. and Somerville,C.
Expressed sequence tags from developing castor seeds
Plant Physiol. 108, 1141-1150 (1995)
CONTACT: Somerville CR
Carnegie Institution
Carnegie Institution
Tel: 4153251521
Email: crs@andrew.stanford.edu
Seq primer: T3.

FEATURES
source Location/Qualifiers
1..405
/organism="Rictinus communis"
/strain="Baker 296"
/db_xref="taxon:3988"
/clone="pcrl345"

/clone.lib="lambdaZAPST"
 /note="Vector: lambdaZAPII, Site_1: EcoRI; Site_2: XhoI;
 Poly(A)+ RNA was purified from developing stage III to
 stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760,
 1982) endosperm plus embryo of immature castor fruits.
 cDNA was synthesized and cloned into lambdaZAPII according
 to the instructions of the manufacturer (Stratagene).
 synthesis was primed from the poly(A) tail, and cloned
 directionally into XhoI (3') and EcoRI (5') sites. In few
 cases, sequence data indicated that this directionality
 was reversed. Partial cDNA clones predominate."
 BASE COUNT 118 a 84 c 86 g 110 t 7 others

Query Match 5.8%; Score 48.8; DB 14; Length 405;
 Best Local Similarity 50.9%; Pred. No. 0.0034; Indels 1; Gaps 1;
 Matches 136; Conservative 0; Mismatches 130;

368 TAGAGAACTCATTCATCCGCTCGGGGCTTCGTTATCCAGCGGCGAGACCCGGGCC 427
 111111 11 111111 111111 111111 111111 111111 111111
 27 TAGAGAGCGCTATCTCAGCGCTTATTTATTTAGTACTTGGGACACGATTCACACTC 86
 111111 11 111111 111111 111111 111111 111111 111111
 428 AAGCTGCTCCCTTAAATCCATTCAGATCTCCGAGCGCGAGATTCATCCCA 487
 111111 111111 111111 111111 111111 111111 111111 111111
 87 TGGCTGCTCTTTATGTTGATTCACAAATGATTCAGAGACGAGATTCACATCA 146
 111111 111111 111111 111111 111111 111111 111111 111111
 488 TCTTTTGGAGGCGCTGCCAATCATTAACAGGGGAGTCATTCTCCGACATGTACA 547
 111111 111111 111111 111111 111111 111111 111111 111111
 147 TTGAGGAGAAATGCGACAGATTAAGTACACCCGAGATCTGACACAGATCTGAG 206
 111111 111111 111111 111111 111111 111111 111111 111111
 548 TGCTGAGCTGGAGACTAGTTGGGCGCAACATTCAC-GCAAGTCCAGACTCTACGAT 606
 111111 111111 111111 111111 111111 111111 111111 111111
 207 TATTACACTTAAAGATATGTTGGGAGACTTCCACTGCAATTCAGAGTCAACCA 266
 111111 111111 111111 111111 111111 111111 111111 111111
 607 GCGGTTTTTAATACCATTTGGGTTG 633
 111111 111111 111111 111111 111111 111111 111111 111111
 267 GGAGCCTTTCGTAAGTCATTCACATG 293
 111111 111111 111111 111111 111111 111111 111111 111111

RESULT 2 479 bp mRNA linear EST 29-NOV-2001
 BG789629 LOCUS
 DEFINITION saes3c07.y1 Gm-c1051 glycine max cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-c1051-8941 5' similar to TR:Q9ZRS1 Q9ZRS1 ALLENE OXIDE
 SYNTHASE: ; mRNA sequence.
 BG789629 ACCESSION
 BG789629.1 GI:14125191
 EST.

soybean.
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 479)
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna,
 A., Bolla, B., Matra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
 Wylie, T., Underwood, K., Steptoe, M., Rheising, B., Allen, M., Bowers,
 Y., Person, B., Swaller, T., Gildous, M., Pape, D., Harvey, N., Schurk,
 R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,
 R., Waterston, R. and Wilson, R.
 TITLE Public Soybean EST Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Shoemaker R/Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 Or contact via email: ccu@resgen.com

FEATURES
 source
 High quality sequence stop: 362.
 Location/Qualifiers
 1..479

/organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1051-8941"
 /clone.lib="Gm-c1051"
 /tissue-type="floral meristematic mRNA"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+, Site_1: EcoRI; Site_2:
 XhoI; The cDNA library was constructed from floral
 meristematic mRNA provided by Dr. Halina Knap of Clemson
 university. Complementary DNA was synthesized from mRNA
 using a primer consisting of a poly(dT) sequence with a
 blunt-ended cDNA fragments followed by XhoI digestion. The
 cDNA fragments were directionally cloned into the
 EcoRI-XhoI restriction site of the pBluescript vector. The
 ligated cDNA fragments were transformed into DH10B host
 cells (GibcoBRL). This library was constructed in the
 laboratory of Dr. Randy Shoemaker."
 BASE COUNT 109 a 187 c 82 g 98 t 3 others

Query Match 5.2%; Score 43.6; DB 12; Length 479;
 Best Local Similarity 55.2%; Pred. No. 0.12; Indels 69; Gaps 0;
 Matches 85; Conservative 0; Mismatches 69;

173 CGCGCGCATGCGCTTACCAATCTGTGCTGCTTACCAAGCGGCGACCAATCTT 232
 111111 111111 111111 111111 111111 111111 111111 111111
 181 CTGGCCCTTCTCCGCCCCCAACCAACGTCGCTTCTCGACGCAAAACCTTCC 240
 111111 111111 111111 111111 111111 111111 111111 111111
 233 ACTTTTGGGAGCGACACAGCGGCGGAAAGCATCTTCCAGCGACACGAT 292
 111111 111111 111111 111111 111111 111111 111111 111111
 241 CCATCTCTTCCACACTCCAAAGCTGACAAAAGAGATGTTCCAGCGACCTTATGC 300
 111111 111111 111111 111111 111111 111111 111111 111111
 293 CCTCTCTCCATTCACCGGAACTTACACAGATCT 326
 111111 111111 111111 111111 111111 111111 111111 111111
 301 CTTCCACCACTCACCGCGGCTACAGATCT 334
 111111 111111 111111 111111 111111 111111 111111 111111

RESULT 3 558 bp mRNA linear EST 06-SEP-2000
 BE660625 LOCUS
 DEFINITION 5-F7 GmaxSC glycine max cDNA, mRNA sequence.
 BE660625 ACCESSION
 BE660625.1 GI:9986517
 EST.

soybean.
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE 1 (bases 1 to 558)
 AUTHORS Harris, N., Chapman, B. P. and Gijzen, M.
 TITLE Gene expression in developing soybean seed coats
 JOURNAL Unpublished (2000)
 COMMENT Contact: Gijzen M
 Agriculture and Agri-Food Canada
 1391 Sandford Street, London, Ontario, Canada N5V 4T3
 Tel: 519 457 1470
 Fax: 519 457 3997
 Email: gijzenm@agr.ca
 Location/Qualifiers
 1..558

/organism="Glycine max"
 /cultivar="Haro63"
 /db_xref="taxon:3847"
 /clone.lib="GmaxSC"
 /tissue-type="Seed coats"
 /lab_host="E. coli strain XL0LR"
 /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI; This

cDNA library was constructed from polyA+ enriched mRNA from green seed coats in mid to late developmental stage, average fresh weight 250 mg per seed. Traces of pod and embryo tissue also present. Complementary DNA was synthesized from mRNA using an XhoI-poly(dT) linker-primer. EcoRI adapters were ligated to the blunt-ended cDNA fragments and the products were digested with XhoI for directional cloning into lambda ZAP Express vector. This lambda library was amplified once using E. coli host strain XL1 Blue MR'. Inserts were then subcloned by mass excision using Exsist helper plasmid for conversion into phagemid vector pBK-CMV in E. coli host strain XL0L.R'.

BASE COUNT 125 a 219 c 90 g 122 t 2 others
ORIGIN

Query Match 5.2%; Score 43.6; DB 10; Length 558;
Best Local Similarity 55.2%; Pred. No. 0.13;
Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

OY 173 CGGCGCATCGACGTTACCATCTGTACGTGGCTTACCAAGCGCAGCAATCCT 232
DB 317 CTGGCCCTCTCTCGGCCCAACCAACGTCGTCTCTCCGACGCAAAACCTTCC 376
OY 233 ACTTTTGGCGACGACGACGCGGGAAGCATCTTTCACCGCAGCAGCAT 292
DB 377 CCATCTCTTGCACACTCCAAAGTTCGACAAAGAGATGTTTTCACCGCAGCCTTCATGC 436
OY 293 CCTCTCTCCCATTCACCGGAGCTACAGATCT 326
DB 437 CTTCCACCACTCACCGCGGCTTACAGATCCT 470

RESULT 4
BE034055 811 bp mRNA linear EST 07-JUN-2000
LOCUS MG04C05 MG Mesembryanthemum crystallinum cDNA 5' similar to
DEFINITION ribosome-inactivating protein gelonin, mRNA sequence.
ACCESSION BE034055
VERSION BE034055.1 GI:8329064
KEYWORDS EST.
SOURCE common iceplant.
ORGANISM Mesembryanthemum crystallinum

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae; Caryophyllales; Alzooceae; Mesembryanthemum.
AUTHORS Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J., Eaton, M., Ferreira, S., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C., Seara, G., Wheeler, M. and Zepeda, G.R.
TITLE Functional Genomics of Plant Stress Tolerance
JOURNAL Unpublished (2000)
COMMENT Contact: Michalowski, C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: cbmeh.arizona.edu
An open reading frame exists.
Insert Length: 1 Std Error: 0.00.

FEATURES
source 1.811
Location/Qualifiers
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_11b="MG"
/tissue_type="roots"
/dev_stage="5-6 weeks"
/note="grown in hydroponics, stress 400 mM NaCl (in 0.5 Hoagland's), 30 h stress"

BASE COUNT 240 a 223 c 162 g 185 t 1 others
ORIGIN

Query Match 5.1%; Score 42.8; DB 10; Length 811;

Best Local Similarity 53.0%; Pred. No. 0.29;
Matches 123; Conservative 0; Mismatches 97; Indels 12; Gaps 1;

OY 110 CCGTCTGGATGCCAAGATTTGTGGTGAACACCAACCAATCAGGGGGAGACTGCA 169
DB 358 CCGGACCGGATCTCCACAGATTCGCTCGTCGACCTCAAAACACCTCCAAAGACCA 417
OY 170 TCAGGGCGGATGAGGATTCACCAATGTAGCTGTGCTTACCAAGC----- 218
DB 418 TCACATCTGCAATGACGTGACAGCCTCTATGCTGGGCTTACCGCAGCAAGCTTGCG 477
OY 219 -AGCGACCAATCTTCTTTTGGCGACGACGACGAGCGCGGGAAGCATCTCTTA 277
DB 478 GAAAGACCGTGGCCAACTCTTGGAGATGCTCCACGCTCGTAGGAACAACCTCTCA 537
OY 278 CCGGACCAACCAATCTCTCTCCATTCACCGGAAGTACAGATCTGGA 329
DB 538 AGGCGCGGAGGTTGGCAACATTCGTTGAGGAGTTACATAGCCCTAGA 589

RESULT 5
AM053634 837 bp mRNA linear EST 20-FEB-2001
LOCUS L30-1401T3 Ice plant Lambda Uni-Zap XR expression library, 30 hours
DEFINITION NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-1401
5', mRNA sequence.
ACCESSION AM053634
VERSION AM053634.1 GI:5916827
KEYWORDS EST.
SOURCE common iceplant.
ORGANISM Mesembryanthemum crystallinum

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae; Caryophyllales; Alzooceae; Mesembryanthemum.
AUTHORS Cushman, J.C.
TITLE An expressed sequence tag database for the common ice plant, Mesembryanthemum crystallinum
JOURNAL Unpublished (1997)
COMMENT Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
Poly(A) tail, 10 nt: 838. .847
PCR Primers
FORWARD: T7
BACKWARD: T3
Insert Length: 847 Std Error: 5.00
Plate: L30-15 Row: A Column: 5
Seq primer: T3
High quality sequence stop: 450.

FEATURES
source 1.837
Location/Qualifiers
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone="L30-1401"
/clone_11b="Ice plant Lambda Uni-Zap XR expression library
/tissue_type="leaf, 30 h 0.4M NaCl"
/dev_stage="Six week old"
/note="vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

BASE COUNT 235 a 193 c 219 g 190 t
ORIGIN

Query Match 5.0%; Score 42.2; DB 10; Length 837;
Best Local Similarity 52.8%; Pred. No. 0.45;
Matches 123; Conservative 0; Mismatches 98; Indels 12; Gaps 1;

OY 110 CCGTTCGATGCGCAAGATTTGTGGTGAACACCAACCAATCAGGGGGAGACTGCA 169

and salt-stressed barley
Unpublished (2002)
Contact: Mark A. Fredricksen
Plant Biology
University of Illinois
1201 W Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: bohnertlab@life.uiuc.edu.

JOURNAL COMMENT

FEATURES
SOURCE
location/Qualifiers
1..913
 /organism="Hordeum vulgare"
 /strain="cv tokak"
 /db_xref="taxon:4513"
 /clone="HB105H11_SK_dbl"
 /clone_1lb="Hg"
 /tissue_type="leaf"
 /dev_stage="3 week old"
 /note="6 and 10 hour drought stress by placing plants on moist paper (75% rel. humidity) in light"

BASE COUNT
161 a 358 c 222 g 170 t 2 others

ORIGIN

Query Match	Best Local Similarity	4.8%	Score 40.4;	DB 14;	Length 913;
Matches 80; Conservative		Pred. No. 1.6;		Mismatches 66;	Indels 0; Gaps 0;

OY 173 CGCGCCGCAATCACCTTACCATCTGTAGTCGTGGCTTACCAGCAGGCGCAATCCT 232
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 255 CCGGCCCTTTACTTGCGGGCGGCAGCCGGCGTGTCGCCGTGCACGCCGAAGACTTCC 314
OY 233 ACTTTTGGCGCAGCACCAGCAGCGCGGGAAAGCAATCTCTTCACCGGCACACAGAT 292
- - - - - ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 CAGTCTCTTTCGACGTCGACAAGTCTGAGAAGAAACCTCTTCACCGGCACTACATGC 374
OY 293 CCTCTCCCATTCACCGGAAGCTAC 318
|||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 375 CCTCCACTCACTACCGGAGGCTTC 400

RESULT 10
LOCUS
BQ765173
DEFINITION
EBrc03_S0006_P09_R.root.. 3 week, waterlogged, cv Optic, EBrc03
ACCESSION
Hordeum vulgare cdna clone EBrc03_S0006_P09 5', mRNA sequence.
BQ765173
VERSION
BQ765173.1 GI:21973660
KEYWORDS
EST.
SOURCE
Hordeum vulgare.
ORGANISM
Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 586)
Hadley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
Ramsay,I., Machray,G., Marshall,D.F.M. and Maugh,R.
Development of Barley Transcriptome Resources
Unpublished (2001)
Contact: Maugh R. Marshall DP
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk.

REFERENCE
AUTHORS

TITLE
JOURNAL COMMENT

FEATURES
SOURCE
1..586
 /organism="Hordeum vulgare"
 /cultivar="Optic"
 /db_xref="taxon:4513"
 /clone="EBrc03_S0006_P09"
 /clone_1lb="root, 3 week, waterlogged, cv Optic, EBrc03"
 /tissue_type="root"
 /dev_stage="3 week"

[illegible]

RESULT	11	463 bp	mRNA	linear	EST 24-July-2000
LOCUS	BE425390				
DEFINITION	BE425390	WHE313_H06_H06Z5	Wheat	unstressed seedling shoot cDNA library	
ACCESSION	BE425390	Triticum aestivum	CDNA clone	WHE313_H06_H06	mRNA sequence.
VERSION	BE425390.1	GI:9423233			
KEYWORDS	EST.				
SOURCE	bread wheat.				
ORGANISM	Triticum aestivum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;				
	1 (bases 1 to 463)				
REFERENCE	Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han				
AUTHORS	P., S., Hsja,K.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.T.,				
	Seaton,C.L. and Tong,J.C.				
TITLE	The structure and function of the expressed portion of the wheat				
	genomes				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Olin Anderson				
	US Department of Agriculture, Agriculture Research Service, Pacific				
	West Area, Western Regional Research Center				
	800 Buchanan Street, Albany, CA 94710, USA				
	Tel: 5105595773				
	Fax: 5105595818				
	Email: andersen@w.usda.gov				
	Sequence have been trimmed to remove vector sequence and low				
	quality sequence with phred score less than 20				
	Seq primer: Stratagene SK primer.				
FEATURES	Location/Qualifiers				
source	1..463				

1. 463
/organism="Triticum aestivum"
/cultivar="Chinese spring"
/db_xref="taxon:4565"
/clone="WHE313_H06_H06"
/clone_lib="Wheat unstressed seedling shoot cDNA library"
/tissue_type="Etiolated shoot"
/dev_stage="Five day old seedling"
/lab_host="E. coli SOLR"
/lab="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized;
germinated and grown aseptically in the dark at room
temperature on filter paper with water, mystatin and
cefotaxime in covered crystallization dishes. Shoots were
harvested. The tissue, total RNA, and poly(A) RNA were

		prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the T7 Close Lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the CD Anderson lab (all other authors)."			
BASE COUNT	68 a	207 c	103 g	85 t	
ORIGIN					
Query Match		4.6%;	Score 38.8;	DB 10;	Length 463;
Best Local Similarity		54.1%;	Pred. No. 3;		
Matches		Conservative	0;	Mismatches 67;	Indels 0; Gaps 0;
Oy	173	CGGCGGCATGACAGCTTACCAATCTGTACTGTCGGCTTACCAMGACAGGCAGCAATCCT	232		
Db	7	CGGGCCCTTCAATGGCGGGCGCACGCCGCGTGTTGGCGTCTCAGCCCAAGAAGCTTCC	66		
Oy	233	ACTTTTTCGGGAGCAGCACCAGACGGCCGGGAAGGCAATCTTTCACGGSCAACACCAGAT	292		
Db	67	CCTGCTCTTTCGACGCTGACACAAGGTCGAGAAAGAAACCTTTCCACGGCACCTTACATGC	126		
Oy	293	CCTCTCCCATTTACCGGAAGCTAC	318		
Db	127	CCTCCACCTCCCTCAACGGCGGCTTC	152		

FEATURES	source	Location/Qualifiers
LOCUS	AJ436193	589 bp mRNA
DEFINITION	AJ436193 S00007 Hordeum vulgare	linear EST 15-MAR-2002
ACCESSION	AJ436193	CDNA clone S0000700017B08F1, mRNA
VERSION	AJ436193.1	
KEYWORDS	GI:19524645	
SOURCE	EST.	
ORGANISM	Hordeum vulgare.	
REFERENCE	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae	
TITLE	; Triticeae; Hordeum.	
JOURNAL	1 (bases 1 to 589)	
COMMENT	Saren,A.-M., Tamskanen,J., Paulin,L. and Schulman,A.H.	
	Barley EST's	
	Unpublished (2002)	
	Contact: Schulman AH	
	Institute of Biotechnology	
	University of Helsinki	
	P.O.Box 56 (Ylikankaari 6A), University of Helsinki FIN-00014,	
	Finland.	
	1.589	
	/organism="Hordeum vulgare"	
	/db_xref="taxon:4513"	
	/clone="S000700017B08F1"	
	/clone_1ib="S00007"	
	/dev_stage="Shoot"	
	/note="2,-3,-4 days after germination"	
BASE COUNT	102 a 237 c 147 g 103 t	
ORIGIN		

	Query Match	Similarity	Score	DB	Length
Best Local	Similarity	54.1%			589;
Matches	79;	Conservative	0;	Mismatches	67;
				Indels	0;
				Gaps	0;
QY	173	CGGCGCCATCGACGTTACCATCTCTACGTGTGGCTTACCAAGAGGCGACCAATCCT	232		
Db	292	CGGGCCCCCTTCATGCGCGCGACCCCGGGGTGTCCCGTGTGCGACGCGCAAGACTTCC	351		
QY	233	ACTTTTGGCGACGCGACGACGAGCGGCGCGAAGGCAATCTTTCACCGGACACCCAGAT	292		
Db	352	CGGTGCTCTTCGACGTCACACCAAGGTGGAAGAAGAACTCTTCACCGGACACTACATGC	411		
QY	293	CCTCTCTCCCATTCACCCGAAGCTAC	318		

Db	Accession	Version	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	412	CCTCCACCTCCTCACC	CGCGGCGCTTC	437					
RESULT 13	AV836389	621 bp	mrna	EST 22-JUN-2001					
LOCUS	AV836389								
DEFINITION	vulgate seedling leaves second leaf stage Hordeum vulgare subsp. vulgare cDNA clone basd1c19, mRNA sequence.								
ACCESSION	AV836389								
VERSION	AV836389.1	GI:14528478							
KEYWORDS	EST.								
SOURCE	Hordeum vulgare subsp. vulgare.								
ORGANISM	Hordeum vulgare subsp. vulgare Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae ; Triticeae; Hordeum. 1 (bases 1 to 621)								
REFERENCE	Sato, K. Barley EST sequencing project in NIG and Okayama Univ Unpublished (2001)								
AUTHORS	Contact: Kazuhiro Sato								
TITLE	Research Institute for Bioreources								
JOURNAL	Okayama University, Barley Germplasm Center Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan Email: kzsato@rib.okayama-u.ac.jp, URL: http://www.rib.okayama-u.ac.jp/barley/ Sato, K., Saitoh, D., Takeda, K., Shint, T. and Kohara, Y. Direct submission: database: http://www.shigen.nig.ac.jp/barley/Barley.html. Location/Qualifiers								
FEATURES	1..621 /organism="Hordeum vulgare subsp. vulgare" /cultivar="Haruna Nijo" /db_xref="taxon:112509" /clone="basd1c19" /clone_lib="K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare seedling leaves second leaf stage" /russu_type="seedling leaves" /dev_stage="second leaf stage" 110 t								
BASE COUNT	109 a	251 c	151 g	110 t					
ORIGIN									
Query Match	4.6%	Score 38.8	DB 10	Length 621					
Best Local Similarity	54.1%	Pred. No. 3.6							
Matches	79	Conservative	0	Mismatches	67	Indels	0	Gaps	0
QY	173	CGCGCGCATCGACGCTTACATCTGTACGTGCTGGCTTACCAAGCAGCGACCAATCCT	232						
Db	280	CGGGCCCTTCATGCGGGCGGCGACCCCGCGTGCCTGACGCCAAGACCTTC	339						
QY	233	ACTTTTGGCGGACGACGACGCGGCGGAAGACATCTTTCACCGGACACCAAGAT	292						
Db	340	CGGTCTCTTCACGCTACCAAGGTCGGAAGAAAGAACCTTCACCGGCGCTTACATGC	399						
QY	293	CCTCTCCCATTCACCGGAAGCTAC	318						
Db	400	CCTCCACCTCCTCACC	CGCGGCGCTTC	425					
RESULT 14	BQ767985	623 bp	mrna	linear	EST 26-JUL-2002				
LOCUS	BQ767985								
DEFINITION	EBR008_S0010_J11_R root, 3 week, drought-stressed, cv Optic, EBR008 Hordeum vulgare cDNA clone EBR008_S0010_J11 5', mRNA sequence.								
ACCESSION	BQ767985								
VERSION	BQ767985.1	GI:21976459							
KEYWORDS	EST.								
SOURCE	Hordeum vulgare.								
ORGANISM	Hordeum vulgare Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae ; Triticeae; Hordeum.								

REFERENCE	1 (bases 1 to 623)				
AUTHORS	Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L., Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.				
TITLE	Development of Barley Transcriptome Resources				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Waugh R, Marshall DF Genome Dynamics/Computational Biology Scottish Crop Research Institute Invergowrie, Dundee, DD2 5DA, Scotland, UK Tel: 00 44 1382 562731 Fax: 00 44 1382 562426 Email: est@scri.sarl.ac.uk.				
FEATURES	Location/Qualifiers				
SOURCE	1..623				
	/organism="Hordeum vulgare"				
	/cultivar="Optic"				
	/db_xref="taxon:4513"				
	/clone="EBro08_S0010.j11"				
	/clone_1lp="root, 3 week, drought-stressed, cv Optic, Ebro08"				
	/tissue_type="root"				
	/dev_stage="3 week"				
	/lab_host="DH10B"				
	/note="Vector: pSPOR1; Site.1: Sal I; Site.2: Not I; Non-normalised library, directionally cloned into pSPOR1. Derived from roots of 3 week old drought stressed barley plants. Developed as part of the barley transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating Gene Function) project."				
BASE COUNT	109 a 249 c 156 g 109 t				
ORIGIN					
Query Match	4.6%	Score 38.8;	DB 14;	Length 623;	
Best Local Similarity	54.1%	Pred. No.3.6;			
Matches	79;	Conservative	0;	Mismatches	67;
				Indels	0;
				Gaps	0;
Qy	173	CGCCGCCATCAGCACTTACCAATCTGTACGTGTGGCTTACCAACGACGACCAATCCT	232		
Db	285	CGGGCCCTTACTATGGCGGGCGACCCCGGTTGTGCTCGACGCCAAGACTTTC	344		
Qy	233	ACTTTTTCGCGACGACGACCAACGACGGCGGAAAGCATCTCTTCACCGGCACCAACAT	292		
Db	345	CGGTGCTCTTTCGACGTCACACCAAGTCGAAAGAACCTTTCACCGGACACTACATGC	404		
Qy	293	CCCTCTCCCATTCACCCGGAAGCTAC	318		
Db	405	CCTCACCTCTCCCTCACCGGGCGCTTC	430		
RESULT 15					
BE060255	626 bp mRNA linear EST 22-OCT-2001				
LOCUS	HVSMEG0011M01f Hordeum vulgare pre-anthesis spike EST library				
DEFINITION	HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone				
	HVSMEG0011M01f, mRNA sequence.				
ACCESSION	BE060255				
VERSION	BE060255.2 GI:13154152				
KEYWORDS	EST.				
SOURCE	Hordeum vulgare.				
ORGANISM	Hordeum vulgare.				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae; Triticeae; Hordeum.				
	1 (bases 1 to 626)				
	Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton, R.D., Close, S.J., Oates, R. and Main, D.				
	Development of a genetically and physically anchored EST resource for barley genomics: Morex pre-anthesis spike cDNA library				
	Unpublished (2001)				
	On Jun 9, 2000 this sequence version replaced gi:8404905.				
	Contact: Wing RA				
	Clemson University Genomics Institute				

Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schmale@iastate.edu

Individual basecall and confidence value were assigned using the
Phred software,
(<http://depts.washington.edu/ventures/collabtr/direct/index.htm#br>
rt). Overall sequence quality assessment and vector trimming were
conducted using the Lucy software (<http://www.tigr.org/softlab/lucy/>).

Lucy parameters were set to ensure an overall trimmed quality of
97.5% or better without any vector fragments in the chosen
high-quality region of each sequence. Low-quality bases between the
poly-T and the high-quality region were replaced with N's to serve
as spacers.

PCR Primers

FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)

BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)

Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).

Location/Qualifiers

1. .550
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="MEST606-C10"
/clone_lib="ISUM5-RN"
/tissue_type="mixed"
/lab_host="DH10B"

/note="vector: pT733PAC; Site_1: EcoRI; Site_2: NotI;
Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG),
Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels
(3, 5, 10, 15, 20, 25, 30, DAP), Adventitious roots (65 DAG
) , Tassel (3-39 cm, 53 and 56 DAG), Immature ear (0.2-3.0
cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated
first ear, ear shank, etiolated seedlings, callus,
Cycloheximide-treated callus, Anaerobic treated seedlings
, NAA (a-Naphthalene acetic acid)-treated seedlings,
kinetin-treated seedlings, ACP
(1-aminocyclopropane-1-carboxylic acid)-treated seedlings,
Brassinolide-treated seedlings, ABA (Abscissic acid
)-treated seedlings, GA (Gibberellic acid)-treated
seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA
molecules were generated as follows. First-strand cDNA was
prepared from oligo-dT selected mRNA by priming with a
NotI oligo-dT primer (5'
AACTGAGAGATTCGGCCCGCCAGCAATTTTCTTTTCTTTT). The
resulting DNA:RNA hybrid was treated with KnaSe H and used
as a template for DNA PolI-catalyzed second strand
synthesis. After the addition of EcoRI adaptors, the
ds-cDNAs were digested with NotI and size-selected. The
resulting molecules were directionally cloned into the
EcoRI and NotI sites of the pT733PAC vector. The library
then went through one round of normalization to Cot value
of 5 based on the methods of Marcelo Bento Soares (Genome
Research 6: 791-806, 1996)."

BASE COUNT 140 a 150 c 125 g 135 t
ORIGIN

Query Match 9.5%; Score 33.8; DB 14; Length 550;

Best Local Similarity 52.1%; Pred. No. 12;

Matches 74; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 150 CCAAAATAGATTCAGATTCGGCGGTAACTTTGGACGNTGAGCAATGTCGGAGCT 209
DB 521 CCGGATCAGACAGATTAAGTGGAGCAGCAAGTTGGACCTGGGATTCATCAGAACT 462
QY 210 GATCTCCAGCTTGGCAGTGTGTTGCAATGCAAGTGGCGCAATTCCTCTCGA 269
DB 461 CGTCTCCAGATTCGGAGCAAGATGATCGAGGCTGCTGAGCAGCCGTCGGCCAGCA 402
QY 270 CCACCTTGGCGCTCTCTA 291
DB 401 CGACATCAACTGTTCTCTGCTA 380

Search completed: March 26, 2003, 19:41:20
Job time : 618.156 secs

QY 193 AGCAATGTTGCGACGTGATC 213
 Db 311 GACAAATGACCTCATGCTGATC 331

RESULT 13
 CNS01XEV/c
 LOCUS
 DEFINITION CNS01XEV 855 bp DNA linear GSS 12-MAY-2000
 Tetraodon nigroviridis genome survey sequence PUC-Or1 end of clone
 202422 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION AL171616
 VERSION AL171616.1 GI:7809673
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 855)
 AUTHORS Roest-Crollius,H., Jallou,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brotlier,P., Queller,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

TITLE Unpublished
 JOURNAL 2 (bases 1 to 855)
 REFERENCE Roest-Crollius,H., Jallou,O., Dasilva,C., Fizames,C., Fisher,C.,
 AUTHORS Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and
 TITLE Weissenbach,J.
 JOURNAL Characterization and repeat analysis of the compact genome of the
 3 (bases 1 to 855)
 REFERENCE Unpublished
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-Apr-2000)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.

FEATURES
 source
 1..855
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="202622"
 /clone_1lb="G"
 /note="Genoscope sequence ID : C0AG202BD11SP1-end :
 PUC-Or1"

BASE COUNT 136 a 146 c 286 g 243 t 44 others

ORIGIN

Query Match 9.5%; Score 34; DB 17; Length 855;
 Best Local Similarity 54.5%; Pred. No. 12;
 Matches 67; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 232 TTGTTGCAATGCAATGGTGCCATTCTCTCGACACACCTTGCCTGCTCCTA 291
 Db 291 TAGTACGTACTACACACATCTCATGATCCCTCATGATCACTTCAGCCATCACTA 232

QY 292 AGGTCCGTCGTGATGCGGCCACACATGTCACTCTCTTTCGAAACCAACCGTGGC 351
 Db 231 CTGCCCCCTCATGATGACCTCATGATGTGCGTAGACGATTCGATGCCAAGGAGGC 172

QY 352 ATC 354
 Db 171 CTC 169

RESULT 14
 BF526014/c

LOCUS BF526014 1078 bp mRNA EST 11-DEC-2000
 DEFINITION 602070191P1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4212886
 5', mRNA sequence.

ACCESSION BF526014
 VERSION BF526014.1 GI:11613375
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1078)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plate: LIA9784 row: f column: 23
 High quality sequence stop: 660.

FEATURES
 source
 1..1078
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4212886"
 /clone_1lb="NCI_CGAP_Brn64"
 /tissue_type="gillbladderoma with EGFR amplification"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: brain; Vector: pCMV-Sport6; Site:1; NotI;
 Site:2; SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.57 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 282 a 296 c 350 g 149 t 1 others

ORIGIN

Query Match 9.5%; Score 34; DB 12; Length 1078;
 Best Local Similarity 56.1%; Pred. No. 14;
 Matches 64; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 191 TGACCAATGTTGCGACGTATCTCCAGCTTGCGATCATGTTGGAATGAGTGGTC 250
 Db 810 TGAACACAGTTCTCGGCGCTCTCGATCCGCTGTGATGCTCGATTCCTCGTGGC 751

QY 251 GGCATCTCTCTCTCTGACACACCTTGCCTGCTCCCTAAGTCCGTGTCG 304
 Db 750 GCTCTCTCGCATCTGATGATCTGCTGCTGTACCATGCACCATGGGG 697

RESULT 15
 B0538836 550 bp mRNA linear EST 12-JUN-2002
 LOCUS B0538836/c
 DEFINITION MEST606-C10.T3 ISUM5-RN Zea mays cDNA clone MEST606-C10 3', mRNA
 sequence.

ACCESSION B0538836
 VERSION B0538836
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 550)
 AUTHORS Wen,T.J., Qiu,F., Guo,L., Ashlock,D.A and Schnable,P.S.
 TITLE Expressed Sequence Tags from B73 Maize: various stages and tissues
 JOURNAL including seedlings treated with a variety of hormones
 COMMENT Unpublished (2001)
 CONTACT: Patrick S. Schnable
 Schnable Laboratory


```

BASE COUNT      117 a      145 c      199 g      125 t
ORIGIN
Query Match      9.6%; Score 34.4; DB 17; Length 586;
Best Local Similarity 60.2%; Pred. No. 7.7;
Matches 56; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 255 ATTCTCTCTCTGACACACCTTGGCCGCTGCTCTAGGTCCGTCGTGAGTCCGAGCCGCA 314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 ATTCAAGCGTGTGTCACCGTGGCGGCGCCCAAGACCATGCTGCTGCTGCCCC 362

QY 315 CGATGTCACCTGCACCTTTTCCGACCCACCGT 347
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 GGTGGCCACCAAGTCATGCTCAGCCCGCGCT 329

RESULT 11
BM624233      632 bp      mRNA      linear      EST 26-FEB-2002
LOCUS
DEFINITION
17000687490749 A.Gam.ad.cDNA1 Anopheles gambiae cDNA clone
ACCESSION
BM624233
VERSION
BM624233.1 GI:18923744
KEYWORDS
EST;
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE
1 (bases 1 to 632)
Holt,R.A., Lin,J.-C., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
Celera Anopheles gambiae EST project
Unpublished (2002)
COMMENT
Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HOLTRA@celera.com
Plate: N0010049WJ row: C column: 17
Seq primer: M13 Reverse.
FEATURES
Source
Location/Qualifiers
1..632
/organism="Anopheles gambiae"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="1960044964539"
/clone_id="A.Gam.ad.cDNA1"
/dev_stage="Adult"
/lab_host="DH10b"
/note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen.
cDNA inserts >500 bp cloned directionally into pSport 1.
Not 1 site is 3'. Clones available through the Malaria
Research and Reference Reagent Resource Center
(www.malaria.mr4.org)."
BASE COUNT      109 a      237 c      169 g      117 t
ORIGIN
Query Match      9.6%; Score 34.4; DB 13; Length 632;
Best Local Similarity 60.2%; Pred. No. 8;
Matches 56; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 255 ATTCTCTCTCTGACACACCTTGGCCGCTGCTCTAGGTCCGTCGTGAGTCCGAGCCGCA 314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 417 ATTCAAGCGTGTGTCACCGTGGCGGCGCCCAAGACCATGCTGCTGCTGCCCC 476

QY 315 CGATGTCACCTGCACCTTTTCCGACCCACCGT 347
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 477 GGTGGCCACCAAGTCATGCTCAGCCCGCGCT 509

```

```

RESULT 12
BG270991      337 bp      mRNA      linear      EST 13-MAR-2002
LOCUS
DEFINITION
BG270991.y1 Melton Mouse Adult Pancreas 1 Mus musculus cDNA clone
IMAGE:5647170 5 similar to SW:TRIP_MOUSE P07146 TRIPSTOEN
PRECUSOR ;, mRNA sequence.
ACCESSION
BG270991
VERSION
BG270991.1 GI:12978684
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 337)
Melton,D., Meadows,A., Clifton,S., Hillier,L., Merra,M., Pape,D.,
Wylie,T., Martin,J., Blistein,A., Schmitt,A., Theising,B., Ritzer
,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole
,R., Tsagarisvill,R., Williams,T., Jackson,Y. and Bowers,Y.
WashU-Harvard Pancreas EST Project
Unpublished (2000)
COMMENT
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@iobhp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
Trace considered overall poor quality
MGI:1853890 This sequence now available from the IMAGE consortium,
for clone orders contact: info@image.llnl.gov
Seq primer: -40RP from Gluco
High quality sequence stop: 1.
FEATURES
Source
Location/Qualifiers
1..337
/organism="Mus musculus"
/strain="TCR"
/db_xref="taxon:10090"
/clone="IMAGE:5647170"
/clone_id="Melton Mouse Adult Pancreas 1"
/sex="Male"
/tissue_type="Total Pancreas"
/dev_stage="Adult"
/lab_host="TOP10"
/note="Organ: Pancreas; Vector: pZERO-2; Site_1: Not I;
Site_2: Xho I; Library constructed using SuperScript
Plasmid library kit (Life Technologies). cDNA made by
oligo-dT priming. Xho I site destroyed during cloning.
Size-selected by column fractionation. Primary library,
unamplified."
BASE COUNT      76 a      86 c      85 g      90 t
ORIGIN
Query Match      9.5%; Score 34; DB 12; Length 337;
Best Local Similarity 50.7%; Pred. No. 7.9;
Matches 102; Conservative 0; Mismatches 98; Indels 1; Gaps 1;

QY 13 CCCATCMTGTGAGGCTTCGCCGGAATTAACAGTGGGAGTCCATCCACCAACATG 72
      || | ||||| ||||| || | ||||| || | ||||| || | ||||| || |
Db 132 CCACCTTGTGTGAGGTTTCCCATCATCATGACAGTGGTGTG- GTCTGCAGCTCAGTGC 190

QY 73 TACATGCTCGACGTCGACAGAGTTGGGTGACATCCACCAAGTCAGCAAGTCCAAG 132
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 191 TACAATATCCGATCCAGATGAGACTGGAGAGACACATCATATGTCCTGGAGGCAAT 250

QY 133 GATGCAATTTTAAATACCAATAGATTGAGATTTCGGCGGAATCTTGTGACGNTG 192
      || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 251 GAGCAGTTTGTGATTGTCACAGATCATCCCGCCCAATTTATTCATGAGACCCCTG 310

```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo. 1 (bases 1 to 879)	NIH-MGC	http://mgc.nci.nih.gov/ .	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
Contact: Robert Strausberg, Ph.D.				

BASE COUNT	284 a	209 c	182 g	204 t
ORIGIN				

RESULT 9					
LOCUS	AJ284561/c				
DEFINITION		485 bp	mRNA	linear	EST 30-JUN-2000
ACCESSION	AJ284561	4A38-P2A9-F	Anopheles gambiae	immune competent	4A38 Anopheles
VERSION	AJ284561	gambiae cDNA clone 4A38-P2A9,			mRNA sequence.
KEYWORDS	AJ284561.1	GI:6932440			
SOURCE	EST.				
	African malaria mosquito.				

FEATURES
SOURCE
 1. 485
 location/qualifiers
 Meyerhofstrasse 1, 69117 Heidelberg, Germany
 European Molecular Biology Laboratory
 Fotis C. Kafatos laboratory
 F. C. Kafatos laboratory

BASE COUNT	95 a	111 c	159 g	120 t
ORIGIN				
Query Match	9.6%	Score 34.4;	DB 9;	Length 485;
Best Local Similarity	60.2%;	Pred. No. 7;		
Matches 56;	Conservative 0;	Mismatches 37;	Indels 0;	Gaps 0.
QY	255	ATTCTCCTCTCTTGACCAACCCCTTCGCCGCTGCTCTTAGGTCGCTGGATGCGGCCAA	314	
Db	437	ATTCAACGCGTGTGTCACACCGTGAGCGCGGTGGCGGCCCAAGACCATCTCGCTGTGCGCC	378	
QY	315	CGATGTCACCTGCACINTTTCGGAACCCACCGT	347	
Db	377	GGTGGCCACCAAGTCATTGCTACACCGCGCGT	345	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	JOURNAL	REFERENCE	AUTHORS	JOURNAL	COMMENT
CNS07HYF/c													
CNS07HYF	Anopheles gambiae GSS T7 end of clone J1J01 of library NotreDame1 from strain PST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.	AL611673	AL611673.1	GI:15963096	GSS.	African malaria mosquito.	1 (bases 1 to 586)	Roth,C.W., Brey,P.T., Ke,Z. and Collins,F.H.					
							Genoscope.						
							Submitted (01-OCT-2001) Genoscope - Centre National de Sequencage :						
							BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr						
							- Web : www.genoscope.cns.fr)						
							2 (bases 1 to 586)						
							Submitted (01-OCT-2001) BBMI, Institut Pasteur, 25, rue du Dr.						
							Roux, Paris 75015, France						
							This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.						

REFERENCE AUTHORS	TITLE	JOURNAL	MEDLINE NUMBER	COMMENT
1 (bases 1 to 485) Dimopoulos, G., Cassavant, T. L., Chang, S., Schneetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M. B. and Katalos, F. C.	Anopheles gambiae pilot gene discovery project: identification of mosquito to innate immunity genes from expressed sequence tags generated from immune-competent cell lines	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)	20300950	Contact: Dimopoulos G

Db 489 CAGCGGGGTCCTTAAGCTTAGAAGTTCTGGGCTCGAGAGATCAGGTGTTTCT 430
 QY 263 CTTCTGACCAACCTTCGCGCTGCTCTTAAGT 295
 Db 429 TTGTGTTGCTGCTTCCTCCCTCTCTCTCTGAGCT 397

RESULT 4

BI474176 554 bp mRNA linear EST 27-AUG-2001
 LOCUS f35h12.y3 zebrafish gridded kidney Danio rerio cDNA clone 4745350
 DEFINITION 5', similar to TR:Q9PVP8 Q9PVP8 CELL-ADHESION PROTEIN ALPHA
 CATEININ, mRNA sequence.

ACCESSION BI474176
 VERSION BI474176
 KEYWORDS GI:15303318
 SOURCE EST

ORGANISM

zebrafish.
 Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
 ; Cyprinidae; Danio.
 1 (bases 1 to 554)

REFERENCE

AUTHORS

Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
 , S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood
 , K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
 Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Ritter, E.,
 Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
 and Wilson, R.

TITLE

Washu zebrafish EST Project 1998
 Unpublished (1998)

JOURNAL

COMMENT

Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrfish@wustl.edu

cDNA library Preparation: Leonard Ira Zon DNA Sequencing by:
 Washington University Genome Sequencing Center Clone distribution:
 Genome Systems, St. Louis, Missouri (web address:
 www.genomesystems.com) (email contact: info@genomesystems.com) and
 Research Genetics, Huntsville, Alabama (web address: www.resgen.com
) (email contact: info@resgen.com) and
 ResourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
 www.rzpd.de)

FEATURES

source

High quality sequence stop: 343.
 Location/Qualifiers
 1..554

/organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone="4745350"
 /clone_lib="zebrafish gridded kidney"
 /sex="mixed"
 /tissue_type="kidney pooled from 300 wild type adults"
 /lab_host="XIOIR"
 /note="Organ: Kidney; Vector: pBR-CMV; Site_1: EcoRI;
 Site_2: XhoI; Oligo dT cDNA library constructed from mRNA
 pooled from pooled kidney tissue from 300 adult
 zebrafish."

BASE COUNT

168 a 122 c 142 g 122 t

Query Match 10.1%; Score 36; DB 13; Length 554;
 Best Local Similarity 53.2%; Pred. No. 2.3;
 Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 214 TCCAGCTTGGCGATCATGTTGTCGATGACGAGTTCCTCTCTGACAC 273
 Db 415 TCCAGCTGAGGCTCCATTCATGAGCTTCGAGCCGCTCTCTTGTATGATGA 356
 QY 274 CCTTCGCGGCTGCTCTTAAGTCCGTCGATGCGGCAAGATGACCTGCACACTMT 333
 Db 355 CCTTTGTTGCTGAGTTTACAGTGTGTGACCTGGGTTACCAATGGCTCAGCAGTCTT 296

QY 334 TCCGAACCCAGCGTGGCATC 354
 Db 295 TCCGACGACAGCGTACGATC 275

RESULT 5

BI885867 749 bp mRNA linear EST 12-OCT-2001
 LOCUS ZF637-1-000122 zebrafish shield stage whole embryo cDNA library
 DEFINITION MPMGP637 Danio rerio cDNA clone MPMGP637_F096 5', mRNA
 sequence.

ACCESSION BI885867
 VERSION BI885867
 KEYWORDS GI:16093138
 SOURCE EST

ORGANISM

zebrafish.
 Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
 ; Cyprinidae; Danio.
 1 (bases 1 to 749)

REFERENCE

AUTHORS

TITLE

EST sequencing of a zebrafish shield stage cDNA library normalised
 by oligonucleotide fingerprinting
 Unpublished (2001)

JOURNAL

COMMENT

Contact: Hennig S
 Laboratory 123, Dept. Lehrnach
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr.63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1612
 Fax: +49 30 8413 1380
 Email: hennig@molgen.mpg.de
 5' EST sequencing of clones from a zebrafish shield stage library,
 normalised from 55,000 starting clones by oligonucleotide
 fingerprinting
 High quality sequence stop: 749.
 Location/Qualifiers
 1..749

FEATURES

source

/organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone="MPMGP637_F096"
 /clone_lib="zebrafish shield stage whole embryo cDNA
 library MPMGP637"
 /tissue_type="whole embryo"
 /dev_stage="shield stage, 6 hrs post-fertilisation"
 /lab_host="E.coli, XL1 blue MR"
 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI;
 oligo-dT-NotI primed, SalI adaptor, directionally cloned,
 library normalised by oligonucleotide fingerprinting"

BASE COUNT 182 a 236 c 150 g 181 t

Query Match 10.1%; Score 36; DB 13; Length 749;
 Best Local Similarity 53.2%; Pred. No. 2.7;
 Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 214 TCCAGCTTGGCGATCATGTTGTCGATGACGAGTTCCTCTCTGACAC 273
 Db 409 TCCAGTGGGCTCCATTCATGAGCTTCGAGCCGCTCTCTTCACTTATGAGGCG 350
 QY 274 CCTTCGCGGCTGCTCTTAAGTCCGTCGATGCGGCAAGATGACCTGCACACTMT 333
 Db 349 CCTTTGTTGCTGAGTTTACAGTGTGTGACCTGGGTTACCAATGGCTCAGCAGTCTT 290
 QY 334 TCCGAACCCAGCGTGGCATC 354
 Db 289 TCCACAGCTAGCGTACGATC 269

RESULT 6

LOCUS

DEFINITION

BG880257 491 bp mRNA linear EST 13-MAR-2002
 BG880257 1b68B04.Y1 Melton Amplified Mouse E10 5 12 5 Pancreas 1 M10S1-A Mus
 musculus cDNA clone IMAGE:5651670 5' similar to TR:Q9R017 Q9R017

Query Match	8.7%;	Score 31;	DB 10;	Length 433;
Best Local Similarity	53.3%;	Pred. No. 0.46;		
Matches	64;	Conservative	0;	Mismatches 56;
			Indels	0;
			Gaps	0;

OY	193	AAGATGTTGGCAGCGCATCTCACGTGGGCACATTGTTTCCGAATCAATGGCGG	252
Dδ	322	ACCANTGATCATCTTGATGGGCCCTCTACCATCATGATGGCATCATGACAAGAACCG	263
OY	253	CCATTCTCTCTCTCGACACACCTTGGCGCGTCTCCTAAGTCGCGTCGATGAGGGCC	312
Dδ	262	TCTCTCTCTCCATAAACTGCTGGCGCGCCGACGAGATGAAGAGTCATGCTTBGCC	203

```

RESULT 15
US-10-015-219-462
: Sequence 462, Application US/10015219
: Publication No. US2003000829A1
: GENERAL INFORMATION:
: APPLICANT: Algate, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
: FILE REFERENCE: 210121.493c1
: CURRENT APPLICATION NUMBER: 2002-03-02
: CURRENT FILING DATE: 2002-03-02
: NUMBER OF SEQ ID NOS: 1739
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 462
: LENGTH: 402
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-015-219-462

```

Query Match	8.68;	Score 30.8;	DB 9;	Length 402;
Best Local Similarity	55.78;	Pred. No. 0.53;		
Matches	59;	Conservative	0;	Mismatches 47;
			Indels	0;
			Gaps	0;

Qy 207 CGAATCTCCAGACCTGGGACATCATGTTGCAATGGAGAGGCTGGGCAATCTCTCTCT 266
 |||||
 Db 94 CTTGATGGCCCTCCTGACGACATCATGATGCAATCATGCAAGGACCCGCTCTGTCTCTCAT 153
 |||||
 Qy 267 CGACCAACCCCTTGGCCGCTGCTCTTAAGTCCCTGCTGATGCGGCC 312
 |||||
 Db 154 AAACGCTCTGGCGCGCCGACGGAATTAAGTGCATGCTCTTGGC 199
 |||||

Search completed: March 26, 2003, 19:45:40
Job time : 57.3643 secs

```

Query Match          8.7%; Score 31.2; DB 9; Length 624;
Best Local Similarity 50.3%; Pred. No. 0.47;
Matches 75; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 140 TTTTAAATACCAATAGATTCGAGATTTCCGCGGTAACTTTGTGAGCNTGAGCAATG 199
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 401 TTTTCTTTAATACATCATCTAATTTGTTTACTTCTTTACGTTAAACATCATTAATCAATG 342
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 200 TTGCGGACGTATCTCCAGCTTGCGGATCATGTTGTCGAATGACATGTCGGCCATTCT 259
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 341 TTATCAACGTCAGCTTTTGGCTTCTTAAAGTCTGTTCAGACCTCTTAATTTGGCTCTCT 282
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 260 CCTCTCGACACACCTTCGCGCGCTGCTC 288
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 281 TCTTTTCAAAATTCCTTTAGACGATCATC 253
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-925-637-51/c
; Sequence 51, Application US/09925637
; Patent No. US2002103338A1
; GENERAL INFORMATION:
; APPLICANT: ChOI
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925,637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-637-51

Query Match          8.7%; Score 31.2; DB 10; Length 624;
Best Local Similarity 50.3%; Pred. No. 0.47;
Matches 75; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 140 TTTTAAATACCAATAGATTCGAGATTTCCGCGGTAACTTTGTGAGCNTGAGCAATG 199
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 401 TTTTCTTTAATACATCATCTAATTTGTTTACTTCTTTACGTTAAACATCATTAATCAATG 342
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 200 TTGCGGACGTATCTCCAGCTTGCGGATCATGTTGTCGAATGACATGTCGGCCATTCT 259
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 341 TTATCAACGTCAGCTTTTGGCTTCTTAAAGTCTGTTCAGACCTCTTAATTTGGCTCTCT 282
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 260 CCTCTCGACACACCTTCGCGCGCTGCTC 288
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 281 TCTTTTCAAAATTCCTTTAGACGATCATC 253
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-08-781-986A-152/c
; Sequence 152, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

```

```

1 STREET: 9410 Key West Avenue
2 CITY: Rockville
3 STATE: Maryland
4 COUNTRY: USA
5 ZIP: 20850
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
9 COMPUTER: HP Vectra 486/33
10 OPERATING SYSTEM: MSDOS version 6.2
11 SOFTWARE: ASCII Text
12
13 CURRENT APPLICATION DATA:
14 APPLICATION NUMBER: US/08/781,986A
15
16 FILING DATE:
17 CLASSIFICATION: 435
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER:
20
21 FILING DATE:
22
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Benson, Bob
25 REGISTRATION NUMBER: 30,446
26 REFERENCE/DOCKET NUMBER: PB248PP
27
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (301) 309-8504
30 TELEFAX: (301) 309-8512
31
32 INFORMATION FOR SEQ ID NO: 152:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 7953 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: double
37
38 TOPOLOGY: linear
39
40 US-08-781-986A-152

```

```

Query Match          8.7%; Score 31.2; DB 7; Length 7953;
Best Local Similarity 50.3%; Pred. No. 1.6;
Matches 75; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY      140  TTTTAAATACCAAAATAGATTCAGATTTCCGCCGGTAACCTTGTGACGNTGAGCAATG 199
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       946  TTTTCTTTTAAATACATCATCTAATTTGTTTACTTTCTTAACTTTAAACATATATTCACATG 887
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      200  TTCCGACAGTGATCTCCAGCTTGCGCATGATGTTTGGAATGACAGTGCGGCATTCCT 259
        || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       886  TTATCAACAGGACATTTTGGCTTCTTAAGTCTGTTCAGACTTCTTAATTTGGGTCTTCT 827
        || | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      260  CCTCTCGACCAACCCCTTCGCCGCGTC 288
        || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB       826  TCTTTTCAAAATTCCTTTAGACGATCATC 798
        || | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-815-343-1076/c
; Sequence 1076, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeline
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon F.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815, 343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1076
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(433)
; OTHER INFORMATION: n = A,T,C or G
; US-09-815-343-1076

```


RESULT 5
US-09-785-921A-3
; Sequence 3, Application US/09785921A
; Patent No. US20020094334A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; FILE REFERENCE: LIT-PI-529
; CURRENT APPLICATION NUMBER: US/09/785,921A
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 3
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Rikinus communis
US-09-785-921A-3

Query Match 14.6%; Score 52.2; DB 10; Length 1731;
Best Local Similarity 49.6%; Pred. No. 2.1e-08;
Matches 129; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATCCTGTGAGAGCTTCGCCGCAAAATTAACAGTGGGAGTCNTCT 60
DB 640 GCAGAGATCCAAATATATTGAGGAGAAATGCCGACGAGATTTAGTACACCGAGATCT 699
QY 61 CCACCAACATGTACATGCTGAGTGGAGAGAGAGAGTGGGTCGCAATCCACCAAGTC 120
DB 700 GCACGAGATCTTACGATTAATACCTTGAGATAGTGGGAGACCTTCAATGCAAT 759
QY 121 CAGCAGTCCAGAGATGATTTTATACCAATATAGATTCGAGTTTCGCCGGTAC 180
DB 760 CAAGAGTCAACAGAGAGAGCTTCTGATTCATTCAGCAAGAGCTAATGTTCC 819
QY 181 TTGTGACGNTGAGCAATGTTGCGAGCTGATCTCCAGCTTGGCAGTATGTTGGAA 240
DB 820 AAATTCAGTGTGATGAGATGATATTAATCCCTATCAGTCTCATGTTATAGA 879
QY 241 TGCAGTGTGCGGCATCTTC 260
DB 880 TCGCAGCTCCACCATCTGC 899

RESULT 6
US-09-785-921A-10
; Sequence 10, Application US/09785921A
; Patent No. US20020094334A1
; GENERAL INFORMATION:
; APPLICANT: Keener, William K.
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: SELECTIVE DESTRUCTION OF CELLS INFECTED WITH HUMAN
; FILE REFERENCE: LIT-PI-529
; CURRENT APPLICATION NUMBER: US/09/785,921A
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 10
; LENGTH: 1805
; TYPE: DNA
; ORGANISM: Rikinus communis
US-09-785-921A-10

Query Match 14.6%; Score 52.2; DB 10; Length 1805;
Best Local Similarity 49.6%; Pred. No. 2.2e-08;
Matches 129; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATCCTGTGAGAGCTTCGCCGCAAAATTAACAGTGGGAGTCNTCT 60
DB 648 GCAGAGTTCAAATATATTGAGGAGAAATGCCGACGAGATTTAGTACAAACGAGATCT 707

QY 61 CCACCAACATGTACATGCTGAGAGAGAGTGGGTCGACAAATCCACCAAGTC 120
DB 708 GCACGAGATCTTACGATTAATTAACCTTGAGATAGTGGGAGAGCTTCAACTCAAT 767
QY 121 CAGCAGTCCAGAGATGATTTTATACCAATATAGTTCACATTTCCGCCGGTAC 180
DB 768 CAAGAGTCTAACAGAGAGAGCTTCTGATTCACAAATTCAGCAAGAGCTAATGTTCC 827
QY 181 TTGTGACGNTGAGCAATGTTGCGAGCTGATCTCCAGCTTGGCAGTATGTTGGAA 240
DB 828 AAATTCAGTGTGATGAGATGATATTAATCCCTATCAGTCTCATGTTATAGA 887
QY 241 TGCAGTGTGCGGCATCTTC 260
DB 888 TCGCAGCTCCACCATCTGC 907

RESULT 7
US-09-815-242-7727
; Sequence 7727, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7727
; LENGTH: 1977
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1977)
US-09-815-242-7727

Query Match 9.1%; Score 32.6; DB 10; Length 1977;
Best Local Similarity 48.1%; Pred. No. 0.25;
Matches 89; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 170 CCGCGGTAACCTTGTGACGNTGAGCAATGTTGCGAGCTGATCTTCACCTTGGGAGTCA 229
DB 881 CTTTCGGTATCTCTCCGAGTGCACCGACCTTCGCCGCAAGGCAATGTTGGCTACA 940
QY 230 TGTGTTGCAATGCAAGTGTGCGCATTTCTCTCTGACCAACCTTGCCTGCTCC 289
DB 941 AGTGATGTGCTGAGCGGCGCGGATGATCACTTCTCGGCTTCACCGTGGCTGCACG 1000
QY 290 TAAGTCCGTGCTGATGCGGCAACGATGTACCTGACTGCTTTTCGAAACCAACGCTGC 349

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 18:49:17 ; Search time 40.3643 Seconds
(without alignments)
7526.383 Million cell updates/sec

Title: US-09-627-165E-13

Perfect score: 357

Sequence: 1 gccagatccaatccatcatt.....aacccacgcgtgcacgtga 357

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCY_NEM_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCYUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185.8	52.0	756	10	US-09-347-064-7 Sequence 7, Appli
2	182.2	51.0	762	10	US-09-347-064-1 Sequence 1, Appli
3	53.8	15.1	1698	10	US-09-785-921A-1 Sequence 1, Appli
4	53.8	15.1	1731	10	US-09-785-921A-2 Sequence 2, Appli
5	52.2	14.6	1731	10	US-09-785-921A-3 Sequence 3, Appli
6	52.2	14.6	1805	10	US-09-785-921A-10 Sequence 10, Appli
7	32.6	9.1	1977	10	US-09-815-242-7727 Sequence 7727, Ap
8	31.6	8.9	2607	10	US-09-391-340-1 Sequence 1, Appli
9	31.6	8.9	2607	10	US-09-948-369-1 Sequence 1, Appli
10	31.4	8.8	15734	10	US-09-764-869-1922 Sequence 1922, Ap
11	31.2	8.7	624	9	US-10-084-205-51 Sequence 51, Appli
12	31.2	8.7	624	10	US-09-925-637-51 Sequence 51, Appli
13	31.2	8.7	7953	7	US-08-781-986A-152 Sequence 152, App
14	31	8.7	433	10	US-09-815-343-1076 Sequence 1076, Ap
15	30.8	8.6	402	9	US-10-015-219-462 Sequence 462, App
16	30.8	8.6	402	10	US-09-777-564-462 Sequence 462, App
17	30.8	8.6	415	9	US-10-015-219-329 Sequence 329, App
18	30.8	8.6	415	10	US-09-777-564-329 Sequence 329, App
19	30.8	8.6	1811	10	US-09-822-830A-617 Sequence 617, App

C	20	30.8	8.6	32816	10	US-09-729-094-3	Sequence 3, Appli
	21	29.8	8.3	66109	10	US-09-880-107-3768	Sequence 3768, Ap
	22	29.8	8.3	198285	10	US-09-880-107-3814	Sequence 3814, Ap
	23	29.6	8.3	31728	9	US-10-114-170-64	Sequence 64, Appl
C	24	29.2	8.2	494	10	US-09-815-343-1378	Sequence 1378, Ap
	25	29.2	8.2	527	10	US-09-070-927A-676	Sequence 676, App
C	26	29.2	8.2	2065	10	US-09-729-920-1	Sequence 1, Appli
	27	28.6	8.0	21636	10	US-09-416-384A-3	Sequence 3, Appli
	28	28.4	8.0	462	9	US-10-046-935-567	Sequence 567, App
	29	28.4	8.0	462	9	US-09-878-178-567	Sequence 567, App
	30	28.4	8.0	494	9	US-10-046-935-16	Sequence 16, Appl
	31	28.4	8.0	494	9	US-09-878-178-16	Sequence 16, Appl
	32	28.4	8.0	495	9	US-10-046-935-124	Sequence 124, App
	33	28.4	8.0	495	9	US-09-878-178-124	Sequence 124, App
	34	28.4	8.0	496	9	US-10-046-935-302	Sequence 302, App
	35	28.4	8.0	496	9	US-10-046-935-603	Sequence 603, App
	36	28.4	8.0	496	9	US-09-878-178-603	Sequence 603, App
	37	28.4	8.0	496	9	US-09-878-178-302	Sequence 302, App
	38	28.4	8.0	496	9	US-09-878-178-603	Sequence 603, App
	39	28.4	8.0	496	9	US-09-878-178-664	Sequence 664, App
C	40	28.4	8.0	896	10	US-09-815-242-3729	Sequence 3729, App
	41	28.4	8.0	1704	10	US-09-815-242-9413	Sequence 9413, Ap
	42	28.4	8.0	2095	9	US-10-025-380-1057	Sequence 1057, Ap
	43	28.4	8.0	2095	10	US-09-922-217-1057	Sequence 1057, Ap
	44	28.4	8.0	2095	10	US-09-833-263-1057	Sequence 1057, Ap
	45	28.2	7.9	6003	12	US-10-010-901-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-347-064-7
; Sequence 7, Application US/09347064A
; Patent No. US20020045208A1
; GENERAL INFORMATION:
; APPLICANT: Eck, Jurgen
; APPLICANT: Schmidt, Arno
; APPLICANT: Zinke, Holger
; TITLE OF INVENTION: Recombinant Fusion Proteins Based on
; TITLE OF INVENTION: Ribosome-Inactivating Proteins of the mistletoe Viscum
; FILE REFERENCE: 09282-5
; CURRENT APPLICATION NUMBER: US/09/347,064A
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: PCT/EP98/00009
; EARLIER FILING DATE: 1998-01-02
; EARLIER APPLICATION NUMBER: EP 97 10 0012.0
; EARLIER FILING DATE: 1997-01-02
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Viscum album
US-09-347-064-7

Query Match 52.0%; Score 185.8; DB 10; Length 756;
Best Local Similarity 82.4%; Pred. No. 1.2e-56;
Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 GCCAGTTCAATCCCTGCTGAGGCTTCCCGCAATTAACAGTGGAGTNCCT 60
|||||
DB 499 GCCAGTTCAATCCCTGCTGAGGCTTCCCGCAATTAACAGTGGAGTNCCT 558
|||||
QY CCACCAATATGATGCTGCTGAGTGGAGTGGGCTGCAATCCCAATGTC 120
|||||
DB 559 CTGCCAGAGTGTACATGCTGAGCTGGAGAGAGATGGGCAATCCACGCAATGTC 618
|||||
QY 121 CAGCAGTCCAGAGTGGATTTTATATACCAATTAAGATTGAGATTTCGCGCGTAAAC 180
|||||
DB 619 CAGCAGTCCAGAGTGGCTTTTATATACCAATTAAGATTGAGATTTCGCGCGTAAAC 678
|||||

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 15:20:00 ; Search time 869.178 Seconds

(without alignments)
11953.474 Million cell updates/sec

Title: us-09-627-165e-13

Perfect score: 357

Sequence: 1 gccacagatcaatccatcctt.....aacccacgcgtgcgcacgtca 357

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Database :

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	353	98.9	357	6	AX139583	AX139583 Sequence
2	230	64.4	572	6	AX139585	AX139585 Sequence
3	225.8	63.2	1873	8	AF369961	AF369961 Viscum al
4	218	61.1	1596	6	AX019438	AX019438 Sequence
5	217.6	61.0	1598	6	AX019435	AX019435 Sequence
6	216.4	60.6	1923	6	AS68957	AS68957 Sequence 1
7	216.4	60.6	1923	6	AR164305	AR164305 Sequence
8	214.8	60.2	1602	6	AY081149	AY081149 Viscum al
9	190.2	53.3	763	6	AX019436	AX019436 Sequence
10	189	52.9	762	6	AX019439	AX019439 Sequence
11	186.4	52.2	747	8	AY081148	AY081148 Viscum al
12	185.8	52.0	705	6	AR164307	AR164307 Sequence
13	185.8	52.0	756	6	A90900	A90900 Sequence 7
14	185.8	52.0	774	6	AR164303	AR164303 Sequence
15	184.2	51.6	768	6	AX139575	AX139575 Sequence
16	182.6	51.1	762	6	AX139571	AX139571 Sequence
17	182.6	51.1	762	6	AF508914	AF508914 Viscum al
18	182.2	51.0	762	6	A90894	A90894 Sequence 1
19	182.2	51.0	762	6	AX139573	AX139573 Sequence
20	179.4	50.3	768	6	AX019440	AX019440 Sequence
21	176.8	49.5	753	8	AF508916	AF508916 Viscum al
22	172.4	48.3	747	8	AF508915	AF508915 Viscum al
23	90	25.2	1596	6	AX019441	AX019441 Sequence
24	78.6	22.0	762	6	AX019442	AX019442 Sequence
25	78.6	22.0	768	6	AX019443	AX019443 Sequence
26	55.4	15.5	1947	6	E01355	E01355 cDNA encodl
27	55.2	15.5	1825	6	AX108095	AX108095 Sequence
28	54.2	15.2	1837	6	AX108130	AX108130 Sequence
29	54.2	15.2	1843	6	AX108123	AX108123 Sequence
30	54.2	15.2	1849	6	AX108116	AX108116 Sequence
31	53.8	15.1	839	6	I07422	I07422 Sequence 34
32	53.8	15.1	839	6	I08354	I08354 Sequence 10
33	53.8	15.1	873	6	A04860	A04860 Artificial
34	53.8	15.1	873	6	A04867	A04867 Artificial
35	53.8	15.1	1050	6	I02107	I02107 Sequence 3
36	53.8	15.1	1050	6	I07418	I07418 Sequence 30
37	53.8	15.1	1050	6	I08353	I08353 Sequence 9
38	53.8	15.1	1140	6	A23903	A23903 partial seq
39	53.8	15.1	1140	6	A23989	A23989 pIC1 1102 p
40	53.8	15.1	1140	6	AR059851	AR059851 Sequence
41	53.8	15.1	1140	6	I25094	I25094 Sequence 15
42	53.8	15.1	1689	8	S40368	S40368 agglutinin
43	53.8	15.1	1695	6	A04859	A04859 Artificial
44	53.8	15.1	1695	6	A10437	A10437 Synthetic n
45	53.8	15.1	1695	6	E00828	E00828 DNA encodln

ALIGNMENTS

RESULT 1
AX139583
LOCUS AX139583 357 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 13 from Patent EP1074560.
ACCESSION AX139583
VERSION AX139583.1 GI:14275213
KEYWORDS
SOURCE Viscum album subsp. coloratum.
ORGANISM Viscum album subsp. coloratum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Santalales; Viscaceae; Viscum.
REFERENCE 1 (bases 1 to 357)
Kim,U., Song,S., Suh,B., Lee,K., Doo,M., Kwak,J., Song,B., Yoon,T.,
AUTHORS Kang,T. and Park,C.

Pred. No. is the number of results predicted by chance to have a

TITLE Crude extract from viscum album coloratum, and proteins and lectins isolated therefrom
JOURNAL Patent: EP 1074560-A 13 07-FEB-2001;
FEATURES Mistle Biotech Co., Ltd. (KR)
SOURCE Location/Qualifiers
 1. 357
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 <1. .>357
CDS /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC39972.1"
 /db_xref="GI:14275214"
 /translation="ARENPIMRLRQINSSESPNNMILETSMGROSTQVOQSK
 DGINFQIRLOISAGNEVYTSNVRDISSLAIMLFECSGREFSSLDHPSPULLNSVVD
 AANDVTCXSEPTVRIY"
BASE COUNT 77 a 103 c 87 g 86 t 4 others
ORIGIN
 Query Match 98.9%; Score 353; DB 6; Length 357;
 Best Local Similarity 100.0%; Pred. No. 1.5e-98;
 Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCAATTCATCCCATGTCGAGGCTTGCGCCGCAATTAACAGTGGGAGTCNTCT 60
 DB 1 GCCAATTCATCCCATGTCGAGGCTTGCGCCGCAATTAACAGTGGGAGTCNTCT 60
 QY 61 CCACCAACATGTACATGCTCGAGCTGAGAGAGAGATTGGGTCGACATCCACCAAGTC 120
 DB 61 CCACCAACATGTACATGCTCGAGCTGAGAGAGAGATTGGGTCGACATCCACCAAGTC 120
 QY 121 CAGCAGTCCAGAGATGGAATTTTAAATACCAATAATATGACATTTCCGCGGTAAAC 180
 DB 121 CAGCAGTCCAGAGATGGAATTTTAAATACCAATAATATGACATTTCCGCGGTAAAC 180
 QY 181 TTTGTAGAGNTGAGCAATGTTTCGAGACGTGATCTCCAGCATTGGGATCATGTTTGGAA 240
 DB 181 TTTGTAGAGNTGAGCAATGTTTCGAGACGTGATCTCCAGCATTGGGATCATGTTTGGAA 240
 QY 241 TGCATGTGTCGGCATTTCTCTCTCTGACCAACCCCTTGCCTCTAAGATCCGTC 300
 DB 241 TGCATGTGTCGGCATTTCTCTCTCTGACCAACCCCTTGCCTCTAAGATCCGTC 300
 QY 301 GTGATGGCGGCAACGATGTACCTGCACCTTTTCCGACACCCGTCGCGATCGTA 357
 DB 301 GTGATGGCGGCAACGATGTACCTGCACCTTTTCCGACACCCGTCGCGATCGTA 357
RESULT 2
AX139585 522 bp DNA linear PAT 30-MAY-2001
LOCUS AX139585
DEFINITION Sequence 15 from Patent EP1074560.
ACCESSION AX139585
VERSION AX139585.1 GI:14275215
KEYWORDS
SOURCE Viscum album subsp. coloratum.
ORGANISM Viscum album subsp. coloratum.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
AUTHORS 1 (bases 1 to 522)
 Kim, J., Song, S., Suh, B., Lee, K., Do, M., Kwak, J., Song, B., Yoon, T.,
 Kang, T., and Park, C.
TITLE Crude extract from viscum album coloratum, and proteins and lectins
 isolated therefrom
JOURNAL Patent: EP 1074560-A 15 07-FEB-2001;
FEATURES Mistle Biotech Co., Ltd. (KR)
SOURCE Location/Qualifiers
 1. 522
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 <1. .>522
 /note="unnamed protein product"
CDS

/codon_start=1
 /protein_id="CAC39973.1"
 /db_xref="GI:14275216"
 /translation="YNDLERYAGHRDPIGLIEELIOSVALRYPGSGTAAQASLLI
 LIQMISEAAEFNDIFWRARQYINSSEFPLDMIMLELETSMGROSTQVOQSDGVENN
 PFRIGISTGNEFTLSNVRDISSLAIMLFVCHDRBRSSDVRWPLVIRVLENSGAVD
 DVCCTASEPTVRIY"
BASE COUNT 112 a 149 c 132 g 129 t
ORIGIN
 Query Match 64.4%; Score 230; DB 6; Length 522;
 Best Local Similarity 78.7%; Pred. No. 2.7e-60;
 Matches 288; Conservative 0; Mismatches 69; Indels 9; Gaps 1;
 QY 1 GCCAATTCATCCCATGTCGAGGCTTGCGCCGCAATTAACAGTGGGAGTCNTCT 60
 DB 157 GCGAATTCATCCCATGTCGAGGCTTGCGCCGCAATTAACAGTGGGAGTCNTCT 216
 QY 61 CCACCAACATGTACATGCTCGAGCTGAGAGAGAGATTGGGTCGACATCCCAAGTC 120
 DB 217 CTTCGCCACATGTACATGCTCGAGCTGAGAGAGAGATTGGGTCGACATCCCAAGTC 276
 QY 121 CAGCAGTCCAGAGATGGAATTTTAAATACCAATAATATGACATTTCCGCGGTAAAC 180
 DB 277 CAGCAGTCCAGAGATGGAATTTTAAATACCAATAATATGACATTTCCGCGGTAAAC 336
 QY 181 TTTGTAGAGNTGAGCAATGTTTCGAGACGTGATCTCCAGCATTGGGATCATGTTTGGAA 240
 DB 337 TTCTGTAGAGNTGAGCAATGTTTCGAGACGTGATCTCCAGCATTGGGATCATGTTTGGAA 396
 QY 241 TGCATGTGTCGGCATTTCTCTCTCTGACCAACCCCTTGCCTCTAAGATCCGTC 300
 DB 397 TGCATGTGTCGGCATTTCTCTCTCTGACCAACCCCTTGCCTCTAAGATCCGTC 456
 QY 301 GTGGA-----TGGCGCAACGATGTACCTGCACCTTTTCCGACACCCGTCGCGC 351
 DB 457 TTGGAATAATAGCGGCGCTGAGAGATTACCTGACATGCTTCCGACACCCGTCGCGC 516
 QY 352 ATCGTA 357
 DB 517 ATCGTA 522
RESULT 3
AF369961 1873 bp DNA linear PLN 12-DEC-2001
LOCUS AF369961
DEFINITION Viscum album subsp. coloratum YCA precursor, gene, complete cds.
ACCESSION AF369961
VERSION AF369961.1 GI:17529700
KEYWORDS
SOURCE Viscum album subsp. coloratum.
ORGANISM Viscum album subsp. coloratum.
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
AUTHORS 1 (bases 1 to 1873)
 Park, W.-B. and Lyu, S.
TITLE Cloning of Viscum album subsp. coloratum (Korean mistletoe)
JOURNAL Biochem. Biophys. Res. Commun. (2002) In press
REFERENCE 2 (bases 1 to 1873)
 Park, W.-B. and Lyu, S.
AUTHORS Direct Submission
TITLE Submitted (13-APR-2001) Chemistry, Seoul Women's University, 126,
 Kongnung-2dong, Nowon-gu, Seoul 139-774, Korea
JOURNAL
FEATURES Location/Qualifiers
 1. 1873
 /organism="Viscum album subsp. coloratum"
 /sub_species="coloratum"
 /db_xref="taxon:159976"
 <138. .>1835
 /product="YCA precursor"
 138. 1835
 /codon_start=1
CDS

```
/product="VCA precursor"  
/protein_id="A140417.1"  
/db_xref="GI:17529701"  
/translation="MALLASRRARVSCFLMLGLVEGATVKAETKERYERLRVYTHO  
ITGDEYFRFTILRDYVSSGSPNEIPILROSTIPVSAORFVVELINOGDSITTA  
IDVYMYVVAOAGDOSYFLRDPADPAERHFTFTGSGYTDLEERYAGHDO  
IPGLIEELIOSVSAIRYSGSTRAORASLLIILOMISAAEPNLFEMARVYINGES  
FLPDYMLELETSMGOSTOYQSDGVDGNNPDRIGTSTGNFVTLNVRDYIASLAIM  
LFVCRDRSSSVRYRPLVIRVLENSAVDVYCTASPEVRIYRGCLVDVDRDK  
FTNGNPIDLPMDPQMLTIRDGTIGSNRGLTYGTYAGYVYVFDCTNAVBKATL  
WOIWNGNITINRSLVLGAASGSSGTLIVQYVFLGQMLAGNDAPREYVITYG  
GNLCMEANGASVYETCGSGSKENRMALYNGSIRPKONODCLISODSSTFENIV  
SCSAGSSORWFETNEGTLINLNLGVMDVADQNSPLRRIIIPATGRKPNQMLPVP"  
138..203  
204..956  
misc_feature  
mat_peptide  
1957..1061  
mat_peptide  
1062..1832  
mat_peptide  
/product="VCA alpha chain"  
/note="Region: linker peptide"  
/product="VCA beta chain"  
BASE COUNT 456 a 458 c 492 g 467 t  
ORIGIN
```

```
Query Match 63.2%; Score 225.8; DB 8; Length 1873;  
Best Local Similarity 78.1%; Pred. No. 6,5e-59;  
Matches 285; Conservative 0; Mismatches 71; Indels 9; Gaps 1;
```

```
QY 1 GCCAGATTCAATCCCATCTGAGGCTTCCGCGCAATTAAAGTGCGGAGTCTCT 60  
DB 735 GCGAGATTCAATCCCATCTTGTGGAGGCTCGCCCAATTAACGCGGAGTCTATT 794  
QY 61 CCACCAACATGTACATGCTCGAGTGAAGAGTGGGTGACATCCACCAAGTC 120  
DB 795 CTTCGCGACATGTACATGCTCGAGTGAAGAGTGGGTGACATCCACCAAGTC 854  
QY 121 CAGCAGTCCAAAGATGGCATTTTAAATACCAATTAAGATGCGATTCGCGGTAAC 180  
DB 855 CACGAGTCTCGAGTGGCGGCTTTTAAATACCAATTCGATGCTATTCACCGGTAAC 914  
QY 181 TTTGTGACGTGAGCAATGTTGCGGACGATCTCCAGCTTGCGGATGCTGCGAA 240  
DB 915 TTGTGACGTGAGCAATGTTGCGGACGATGCTCCAGCTTACGATGTTGTGTA 974  
QY 241 TGCAGTGTGCGCATCTCTCTCGACACCTTCGCGGCTGCTCTAAGTCCGTC 300  
DB 975 TGTAGGAGACGACCATCTCTCGACGATGCTGATTTGGCGCTGTGATACGACCGTC 1034  
QY 301 GTGGA-----TGGGCGCAACGATGCTGACCTGCTTTTCCGAACCCACGTCGC 351  
DB 1035 TTGGAATAATAGCGCGCGCTGACGATGTTACTGCTGACGATCCGAACCCACGTCGC 1094  
QY 352 ATGCT 356  
DB 1095 ATTGT 1099
```

```
RESULT 4  
AX019438 1596 bp DNA linear PAT 07-SEP-2000  
LOCUS AX019438  
DEFINITION Sequence 15 from Patent WO9940109.  
ACCESSION AX019438  
VERSION AX019438.1 GI:10043377  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial sequences.  
REFERENCE  
1 (bases 1 to 1596)  
Welters,P., Stiefel,T., Voelter,W. and Morris,P.  
TITLE  
Recombinant mistletoe lectines  
Patent: WO 9940109-A 15 12-AUG-1999;  
JOURNAL  
WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH  
(DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB)  
FEATURES  
Location/Qualifiers
```

```
source 1..1596  
/organism="synthetic construct"  
/db_xref="taxon:32630"  
BASE COUNT 387 a 418 c 419 g 372 t  
ORIGIN
```

```
Query Match 61.1%; Score 218; DB 6; Length 1596;  
Best Local Similarity 74.7%; Pred. No. 1,7e-56;  
Matches 266; Conservative 13; Mismatches 74; Indels 6; Gaps 1;
```

```
QY 1 GCCAGATTCAATCCCATCTGAGGCTTCCGCGCAATTAAAGTGCGGAGTCTCT 60  
DB 499 GCGAGATTCAATCCCATCTTATGAGGCTCGCCAAATTAACAGTGGCGCTCATTT 558  
QY 61 CCACCAACATGTACATGCTCGAGTGAAGAGTGGGTGACATCCACCAAGTC 120  
DB 559 CTTCGCGACGATGTACATGCTCGAGTGAAGAGTGGGTGACATCCACCAAGTC 618  
QY 121 CAGCAGTCCAAAGATGGCATTTTAAATACCAATTAAGATGCGATTCGCGGTAAC 180  
DB 619 CAGCAGTCCAAAGATGGCATTTTAAATACCAATTAAGATGCGATTCGCGGTAAC 678  
QY 181 TTTGTGACGTGAGCAATGTTGCGGACGATCTCCAGCTTGCGGATCATGTTGTGAA 240  
DB 679 TTGTGACGTGAGCAATGTTGCGGACGATCTCCAGCTTGCGGATCATGTTGTGTA 738  
QY 241 TGCAGTGTGCGCATCTCTCTCGACACCTTCGCGGCTGCTCTAAGTCCGTC 300  
DB 735 TGCAGTGTGCGCATCTCTCTCGACACCTTCGCGGCTGCTCTAAGTCCGTC 798  
QY 301 GTGATGCGGCGCCCAACGATGCTGACCTGCTTTTCCGAACCCACCGCGCATGCT 356  
DB 799 AT-----AGCGAGATGATTACTGCGAGTGTGGAACCTACGATGCTGATTTGT 848
```

```
RESULT 5  
AX019435 1598 bp DNA linear PAT 07-SEP-2000  
LOCUS AX019435  
DEFINITION Sequence 12 from Patent WO9940109.  
ACCESSION AX019435  
VERSION AX019435.1 GI:10043374  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial sequences.  
REFERENCE  
1 (bases 1 to 1598)  
Welters,P., Stiefel,T., Voelter,W. and Morris,P.  
TITLE  
Recombinant mistletoe lectines  
Patent: WO 9940109-A 12 12-AUG-1999;  
JOURNAL  
WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH  
(DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB)  
FEATURES  
source 1..1598  
/organism="synthetic construct"  
/db_xref="taxon:32630"  
misc_feature 319  
misc_feature 1322  
BASE COUNT 373 a 396 c 392 g 356 t 81 others  
ORIGIN
```

```
Query Match 61.0%; Score 217.6; DB 6; Length 1598;  
Best Local Similarity 74.7%; Pred. No. 2,2e-56;  
Matches 266; Conservative 13; Mismatches 71; Indels 6; Gaps 1;
```

```
QY 1 GCCAGATTCAATCCCATCTGAGGCTTCCGCGCAATTAAAGTGCGGAGTCTCT 60  
DB 500 GCGAGATTCAATCCCATCTTATGAGGCTCGCCAAATTAACAGTGGCGMTCATTT 559  
QY 61 CCACCAACATGTACATGCTCGAGTGAAGAGTGGGTGACATCCACCAAGTC 120  
DB 560 CTTCGCGACGATGTACATGCTCGAGTGAAGAGTGGGTGACATCCACCAAGTC 619  
QY 121 CAGCAGTCCAAAGATGGCATTTTAAATACCAATTAAGATGCGATTCGCGGTAAC 180
```

```

Db 620 CAGCATTCAACCGATGCTTTTATAACCCAMTYCGGTGGCTATATACYMCYGTAC 679
QY 181 TTGTGACATGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
Db 680 TTTCGTACCTTMCYTAAGTTGCKMYGTGATCGCAGCTTGGCATATGTTTGTATA 739
QY 241 TGCAGTGTGCGGCATTTCTCTCTCGACACCCCTTGCCTGCTCTTAAGTCCGTC 300
Db 740 TCCGAGAGACGGGCATCTCTCTGACGTCGCTATTTGGCCGCTGCTATACGACCCGTG 799
QY 301 GTGATGCGGCCCAAGATGTACCTCGACTNTTTCGAACCCAGCTGCGCATGCT 356
Db 800 AT-----AGCCGATGATGTTACCTGCACTGCTCGAACCACCTGCGGATGT 849

RESULT 6
AS8957
LOCUS AS8957 1923 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 1 from Patent WO9701636.
ACCESSION AS8957
VERSION AS8957.1 GI:3714428
KEYWORDS
SOURCE
ORGANISM
. Viscum album.
. Viscum album
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Santalales; Viscaceae; Viscum.
REFERENCE
1 (bases 1 to 1923)
Lentzen, H., Eck, J., Baur, A. and Zinke, H.
Recombinant mistletoe lectin (rML)
Patent: WO 9701636-A 16-JAN-1997;
MADAUS AG KOELN (DE)
FEATURES
source
1..1923
location/Qualifiers
/organism="Viscum album"
/db_xref="taxon:3972"
/ctssue_type="LEAF"
55..1749
/codon_start=1
/product="PREPROMISTLETOE-LECTIN"
/protein_id="CAA03513.1"
/db_xref="GI:3714428"
/translaction="MNAVMDSSRRAMASCFIMLGLVFGATVKAETKESYERLRURVTHO
TTGDEYRFITLADYVSSGSFSEIPLDQSTIPVSDARFVLEVLINGGDSITAA
IDVNLVYVYAGDQSYFLDAPRGAEHLFTPTSSSPFGSGPDLDERYGHDO
IPGLIDLIOSVTLALRPGGSTRQASIIILIMTISEARFPIIMRANQYINSGAS
FLPPVWLELETSMGOSTOVHSTGCTGFENPIRLAIPENFYTLNVDYASLAIM
LFNGERSSSEYRMYPLVIRPIADYVTSASEPVRIVGRNGCVDRDDFERDGN
QIOLMPKSNNDPQMTIRDGTIRNSGSLTYTGTAIVYIMIDCNVAVEATIM
QIWMNGTIIINPRSNLVLAASSGIRGTLTVQTLDTLGGQMLAGNDTAPREVTIYGER
DLCMESNGSVMWETCVSQKNOMALYGGGIRPKONODGCTGCRDGSYTVINVS
CSAGSGORWVFNEGAILNLKNGLAMDVANQPKLRIIITPATKRPQMLPVP"
55..153
sig_peptide
mat_peptide
154..900
/product="A-CHAIN"
mat_peptide
901..957
/product="PROPEPTIDE"
mat_peptide
958..1746
/product="B-CHAIN"
BASE COUNT 465 a 488 c 495 g 475 t
ORIGIN
Query Match 60.6%; Score 216.4; DB 6; Length 1923;
Best Local Similarity 77.2%; Pred. No. 5,4e-56;
Matches 275; Conservative 0; Mismatches 75; Indels 6; Gaps 1;
QY 1 GCCAGATTCAATCCCATCCTGTGGAGGCTTCGCCGCAAAATTAACAGTGGGAGTCTCT 60
Db 652 GCCAGATTCAATCCCATCTTATGAGGCTTCGCCCAATTAACAGTGGGAGTCTATTT 711
QY 61 CCACCAAAACATGTACATGCTCGAGCTGGAGAGAGTTGGGGTGCACATCCACCAAGTC 120
Db 712 CTGCCAGAGCTGTACATGCTGAGCTGGAGAGAGTTGGGGCAACCAATCCACGCAAGTC 771
QY 121 CACACATCCCAAGATGGCATTTTAAATACCCAATTAAGATTGACATTTCCCGCGTAA 180
Db 772 CACACATCCCAAGATGGCATTTTAAATACCCAATTAAGATTGACATTTCCCGCGTAA 180
QY 772 CACACATCCCAAGATGGCATTTTAAATACCCAATTAAGATTGACATTTCCCGCGTAA 831
Db 181 TTGTGACGNTGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
QY 832 TTGTGACGNTGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
Db 832 TTGTGACGNTGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
QY 241 TGCAGTGTGCGGCATTTCTCTCTCGACACCCCTTGCCTGCTCTTAAGTCCGTC 300
Db 892 TCCGAGAGAGCGGCATCTCTCTGACGTCGCTATTTGGCCGCTGCTATACGACCCGTG 951
QY 301 GTGATGCGGCCCAAGATGTACCTCGACTNTTTCGAACCCAGCTGCGCATGCT 356
Db 952 AT-----AGCCGATGATGTTACCTGCACTGCTCGAACCACCTGCGGATGT 1001

RESULT 7
AR164305
LOCUS AR164305 1923 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 34 from patent US 6271368.
ACCESSION AR164305
VERSION AR164305.1 GI:16235409
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1923)
Lentzen, H., Eck, J., Baur, A. and Zinke, H.
Recombinant mistletoe lectin (rML)
Patent: US 6271368-A 34-OCT-2001;
MADAUS AG KOELN (DE)
FEATURES
source
1..1923
location/Qualifiers
/organism="unknown"
BASE COUNT 465 a 488 c 495 g 475 t
ORIGIN
Query Match 60.6%; Score 216.4; DB 6; Length 1923;
Best Local Similarity 77.2%; Pred. No. 5,4e-56;
Matches 275; Conservative 0; Mismatches 75; Indels 6; Gaps 1;
QY 1 GCCAGATTCAATCCCATCCTGTGGAGGCTTCGCCGCAAAATTAACAGTGGGAGTCTCT 60
Db 652 GCCAGATTCAATCCCATCTTATGAGGCTTCGCCCAATTAACAGTGGGAGTCTATTT 711
QY 61 CCACCAAAACATGTACATGCTCGAGCTGGAGAGAGTTGGGGTGCACATCCACCAAGTC 120
Db 712 CTGCCAGAGCTGTACATGCTGAGCTGGAGAGAGTTGGGGCAACCAATCCACGCAAGTC 771
QY 121 CACACATCCCAAGATGGCATTTTAAATACCCAATTAAGATTGACATTTCCCGCGTAA 180
Db 772 CACACATCCCAAGATGGCATTTTAAATACCCAATTAAGATTGACATTTCCCGCGTAA 180
QY 772 CACACATCCCAAGATGGCATTTTAAATACCCAATTAAGATTGACATTTCCCGCGTAA 831
Db 181 TTGTGACGNTGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
QY 832 TTGTGACGNTGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
Db 832 TTGTGACGNTGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
QY 241 TGCAGTGTGCGGCATTTCTCTCTCGACACCCCTTGCCTGCTCTTAAGTCCGTC 300
Db 892 TCCGAGAGAGCGGCATCTCTCTGACGTCGCTATTTGGCCGCTGCTATACGACCCGTG 951
QY 301 GTGATGCGGCCCAAGATGTACCTCGACTNTTTCGAACCCAGCTGCGCATGCT 356
Db 952 AT-----AGCCGATGATGTTACCTGCACTGCTCGAACCACCTGCGGATGT 1001

```

```

Db 712 CTGCCAGAGCTGTACATGCTGAGCTGGAGAGAGTTGGGGCAACCAATCCACCAAGTC 771
QY 121 CACAGATCCCAAGATGGCATTTTAAATACCCAATTAAGATTGCAATTCGCCGCTAAC 180
Db 772 CACAGATCCCAAGATGGCATTTTAAATACCCAATTAAGATTGCAATTCGCCGCTAAC 831
QY 181 TTGTGACGNTGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
Db 832 TTGTGACGNTGAGCAATGTTCGCAGCGATCTCCAGCTTGGCATATGTTTCCAA 240
QY 241 TGCAGTGTGCGGCATTTCTCTCTCGACACCCCTTGCCTGCTCTTAAGTCCGTC 300
Db 892 TCCGAGAGAGCGGCATCTCTCTGACGTCGCTATTTGGCCGCTGCTATACGACCCGTG 951
QY 301 GTGATGCGGCCCAAGATGTACCTCGACTNTTTCGAACCCAGCTGCGCATGCT 356
Db 952 AT-----AGCCGATGATGTTACCTGCACTGCTCGAACCACCTGCGGATGT 1001

RESULT 8
AY081149
LOCUS AY081149 1602 bp mRNA linear PLN 17-MAR-2002
DEFINITION Viscum album lectin chain A isoform 1 precursor, mRNA, partial cds.
ACCESSION AY081149

```


VERSION AY081149.1 GI:19526600
KEYWORDS Viscum album.
SOURCE Viscum album
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Santalales; Viscaceae; Viscum.
REFERENCE 1 (bases 1 to 1602)
AUTHORS Paramasivam, M., Misra, V., Sriivasan, A. and Singh, T.P.
TITLE Viscum album (Indian) mRNA for Mistletoe lectin chain A isoform 1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1602)
AUTHORS Paramasivam, M., Misra, V., Sriivasan, A. and Singh, T.P.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Biophysics, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India

FEATURES
source
1..1602
/organism="Viscum album"
/db_xref="taxon:3972"
/rissue_type="leaf"
/country="India"
<1..1596
/note="toxic component"
/codon_start=1
/product="lectin chain A isoform 1 precursor"
/protein_id="AAL87006.1"
/db_xref="GI:19526601"

CDS
misc_feature
748..804
/note="toxic component; ribosome-inactivating protein"
mat_peptide
805..1593
/product="lectin B chain"
3'UTR
1597..1602
BASE COUNT 394 a 415 c 419 g 374 t
ORIGIN
Query Match 60.2%; Score 214.8; DB 8; Length 1602;
Best Local Similarity 77.0%; Pred. No. 1.7e-55;
Matches 274; Conservative 0; Mismatches 76; Indels 6; Gaps 1;

1 GCCAGATTCAATCCCATCTGTGAGGCTCGCGGCAATTACAGTGGAGTCTCT 60
|||||
Db 499 GCCAGATTCAATCCCATCTGTGAGGCTCGCGGCAATTACAGTGGAGTCTCT 558
|||||
QY 61 CCACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
|||||
Db 559 CTGGCAGACGTGTACATCTGTGAGGCTCGCGGCAATTACAGTGGAGTCTCT 618
|||||
QY 121 CAGCAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
|||||
Db 619 CAGCAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 678
|||||
QY 181 TTGTGAGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
|||||
Db 679 TTGTGAGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 738
|||||
QY 241 TGCAGTGGTGGCATTCT 300
|||||
Db 739 TGCAGTGGTGGCATTCT 798
|||||

QY 301 GTGGATCGGCCCAACGATGTCACCTGACATNTTTCGGAACCAACCGTGGCATCGT 356
|
Db 799 AT-----ACCGATGATGTTACCTGACAGTCTTGGGAACCTACGGTGGCATTTGT 848
|||||

RESULT 9
LOCUS AX019436 763 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 13 from Patent WO9940109.
ACCESSION AX019436
VERSION AX019436.1 GI:10043375
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 763)
AUTHORS Welters, P., Stiefel, T., Voelker, W. and Morris, P.
TITLE Recombinant mistletoe lectines
JOURNAL Patent: WO 9940109-A 13 12-AUG-1999;
WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB)

FEATURES
source
1..763
/organism="synthetic construct"
/db_xref="taxon:32630"

misc_feature 163 a 215 c 175 g 170 t 40 others
BASE COUNT
ORIGIN
Query Match 53.3%; Score 190.2; DB 6; Length 763;
Best Local Similarity 78.4%; Pred. No. 6.6e-48;
Matches 207; Conservative 13; Mismatches 44; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCTGTGAGGCTCGCGGCAATTACAGTGGAGTCTCT 60
|||||
Db 500 GCCAGATTCAATCCCATCTGTGAGGCTCGCGGCAATTACAGTGGAGTCTCT 559
|||||
QY 61 CCACCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
|||||
Db 560 CTGGCAGACGTGTACATCTGTGAGGCTCGCGGCAATTACAGTGGAGTCTCT 619
|||||
QY 121 CAGCAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
|||||
Db 620 CAGCAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 679
|||||
QY 181 TTGTGAGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
|||||
Db 680 TTGTGAGAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739
|||||
QY 241 TGCAGTGGTGGCATTCT 264
|||||
Db 740 TGCAGTGGTGGCATTCT 763
|||||

RESULT 10
LOCUS AX019439 762 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 16 from Patent WO9940109.
ACCESSION AX019439
VERSION AX019439.1 GI:10043378
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 762)
AUTHORS Welters, P., Stiefel, T., Voelker, W. and Morris, P.
TITLE Recombinant mistletoe lectines
JOURNAL Patent: WO 9940109-A 16 12-AUG-1999;
WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB)

FEATURES
source
1..762
/organism="synthetic construct"
/db_xref="taxon:32630"

DEFINITION Sequence 7 from Patent WO9829540.
 ACCESSION A90900
 VERSION A90900.1 GI:6739438
 KEYWORDS
 SOURCE European mistletoe.
 ORGANISM
 Viscum album
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 1 (bases 1 to 756)
 AUTHORS Eck, J. and Schmidt, A.
 TITLE RECOMBINANT FUSION PROTEINS BASED ON RIBOSOME-INACTIVATING PROTEINS
 JOURNAL OF EUROPEAN MISTLETOE VISCUM ALBUM
 Patent: WO 9829540-A 7 09-JUL-1998;
 ECK JUERGEN (DE); SCHMIDT ARNO (DE)
 FEATURES
 source
 1..756
 /organism="Viscum album"
 /db_xref="taxon:3972"
 /tissue="LEAF"
 1..756
 /gene="MLA"
 <1..>756
 /gene="MLA"
 /function="TOXIC COMPONENT"
 /note="FIG. 11.A: NUKLEOTIDSEQUENZ UND ABGELEITETE AMINO
 ACIDSEQUENZ VON RMLA"
 /codon_start=1
 /product="RECOMBINANT MISTLETOE LECTIN A-CHAIN."
 /protein_id="CAB69396.1"
 /db_xref="GI:6739439"
 /translation="VERIRLRVHTQTEGEYFRITTLRDYVSSGSFNSPEPLRST
 IPVSDAQRFLVLELTNOGDSITPAIDVNLVVAAYAGDSIFLRAPKAEHTHT
 GTTSSLPFNISYDLPRIYAGHDOIPLQIDQLIQSVTALFPGSGSTGTQARSTLLLI
 OMISEAARFNPIIMRROYINSASFLPDYVMELETSMWGSQGYOVHSTDGVEFNPI
 RLAIIPGNFVLTIVRDYIASLAIIMLFVCGRPS"
 BASE COUNT 167 a 229 c 181 g 179 t
 ORIGIN
 Query Match 52.0%; Score 185.8; DB 6; Length 756;
 Best Local Similarity 82.4%; Pred. No. 1.5e-46;
 Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 1 GCCAGATTCAATCCCATCTGTCGAGCGTTCGCCGCAATTAACAGTGGAGTCT 60
 |||||||
 DB 499 GCCAGATTCAATCCCATCTTATGGAGGCTCGCCCAATTAACAGTGGCGTCATT 558
 |||||||
 QY 61 CCACCAACATGTACATGCTGCGAGCGAGAGAGAGTGGGTCGACCAATCCACCAAGTC 120
 |||||
 DB 559 CTGCCAGAGCTGTACATGCTGAGAGCGAGAGAGAGTGGGCGCAACATCCACGCAAGTC 618
 |||||
 QY 121 CAGCAGTCCAGAGATGCAATTTTAAATACCAATTAAGTTCGACATTTCCCGCGGTAA 180
 |||||
 DB 619 CAGCATTCAACCGATGCGCTTTTAAATACCAATTCGCTTGGCTATACCCCCGGGTAA 678
 |||||
 QY 181 TTGTGACGCTGACATGTCGCGAGCGATCTCCAGTTCGGCGATCATGTTGTGAA 240
 |||||
 DB 679 TTGTGACGCTGACATGTCGCGAGCGATGCGCAGCTGCGCATCATGTTGTTGTA 738
 |||||
 QY 241 TGCAGTGTGCGCCAT 256
 |||||
 DB 739 TGGGAGAGCGGCCAT 754
 |||||
 RESULT 14
 ARI64303 774 bp DNA linear PAT 17-OCT-2001
 LOCUS
 DEFINITION Sequence 30 from patent US 6271368.
 ACCESSION ARI64303
 VERSION ARI64303.1 GI:16235405
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

Unclassified.
 1 (bases 1 to 774)
 AUTHORS Lentzen, H., Eck, J., Baur, A. and Zinke, H.
 TITLE Recombinant mistletoe lectin (rML)
 JOURNAL Patent: US 6271368-A 30 07-AUG-2001;
 FEATURES
 source
 1..774
 /organism="unknown"
 BASE COUNT 173 a 232 c 185 g 184 t
 ORIGIN
 Query Match 52.0%; Score 185.8; DB 6; Length 774;
 Best Local Similarity 82.4%; Pred. No. 1.5e-46;
 Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 1 GCCAGATTCAATCCCATCTGTCGAGCGTTCGCCGCAATTAACAGTGGAGTCT 60
 |||||||
 DB 505 GCCAGATTCAATCCCATCTTATGGAGGCTCGCCCAATTAACAGTGGCGTCATT 564
 |||||||
 QY 61 CCACCAACATGTACATGCTGCGAGCGAGAGAGAGTGGGTCGACCAATCCACCAAGTC 120
 |||||
 DB 565 CTGCCAGAGCTGTACATGCTGAGAGCGAGAGAGAGTGGGCGCAACATCCACGCAAGTC 624
 |||||
 QY 121 CAGCAGTCCAGAGATGCAATTTTAAATACCAATTAAGTTCGACATTTCCCGCGGTAA 180
 |||||
 DB 625 CAGCATTCAACCGATGCGCTTTTAAATACCAATTCGCTTATACCCCCGGGTAA 684
 |||||
 QY 181 TTGTGACGCTGACATGTCGCGAGCGATCTCCAGTTCGGCGATCATGTTGTGAA 240
 |||||
 DB 685 TTGTGACGCTGACATGTCGCGAGCGATGCGCAGCTTGGCGATCATGTTGTTGA 744
 |||||
 QY 241 TGCAGTGTGCGCCAT 256
 |||||
 DB 745 TGGGAGAGCGGCCAT 760
 |||||
 RESULT 15
 AX139575 768 bp DNA linear PAT 30-MAY-2001
 LOCUS
 DEFINITION Sequence 5 from Patent EP1074560.
 ACCESSION AX139575
 VERSION AX139575.1 GI:14275209
 KEYWORDS
 SOURCE Viscum album subsp. coloratum.
 ORGANISM Viscum album subsp. coloratum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Santalales; Viscaceae; Viscum.
 1 (bases 1 to 768)
 AUTHORS Kim, J., Song, S., Suh, B., Lee, K., Do, M., Kwak, J., Song, B., Yoon, T.,
 Kang, T. and Park, C.
 TITLE Crude extract from Viscum album coloratum, and proteins and lectins
 JOURNAL Isolated therefrom
 Patent: EP 1074560-A 5 07-FEB-2001;
 Mistle Biotech Co., Ltd. (KR)
 FEATURES
 source
 1..768
 /organism="Viscum album subsp. coloratum"
 /db_xref="taxon:159976"
 BASE COUNT 171 a 219 c 200 g 178 t
 ORIGIN
 Query Match 51.6%; Score 184.2; DB 6; Length 768;
 Best Local Similarity 80.7%; Pred. No. 4.8e-46;
 Matches 213; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 1 GCCAGATTCAATCCCATCTGTCGAGCGTTCGCCGCAATTAACAGTGGAGTCT 60
 |||||||
 DB 505 GCCAGATTCAATCCCATCTTATGGAGGCTCGCCCAATTAACAGTGGGTCATT 564
 |||||||
 QY 61 CCACCAACATGTACATGCTGAGCGAGAGAGAGTGGGTCGACCAATCCACCAAGTC 120
 |||||
 DB 565 CTGCCAGAGCTGTACATGCTGAGAGCGAGAGAGTGGGCGCAACATCCACCAAGTC 624
 |||||

```

QY 121 CAGCAGTCCAGAGATGGCATTTTATATACCAATATAGATTGCAGATTTCCGCCGGTAA 180
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 625 CAGCAGTGCAGCCGATGGGGTTTATATACCAATTCGGTTGGTATATCCACCGGCAAC 684
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 TTTTGACGNTGAGCAATGTTGCGAGCGTATCTCCAGCTTGGCGATCATGTGTTGGA 240
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 685 TTGTTGTTGAGCAATGTTGCGAGCGTATCCGACCTTGGGATCATGTGTTGTA 744
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 TGCAGTGGTGGCCATTCCTCT 264
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 745 TGCAGGAGCGGTCTCTCCCT 768
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: March 26, 2003, 19:07:00
 Job time : 877.678 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 14:11:30 ; Search time 98.6683 Seconds
(without alignments)
8148.139 Million cell updates/sec

Title: US-09-627-165E-13

Perfect score: 357
Sequence: 1 gccagattcaatccatcctt.....aaccacccgtcgatcgta 357

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: N.Geneseq_101002.*
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353	98.9	357	22 AAC85478	Korean mistletoe 1
2	230	64.4	522	22 AAC85479	Korean mistletoe 1
3	227.4	63.7	1566	24 ABL56947	Galactose recognis
4	218	61.1	1596	20 AAZ09103	Mistletoe lectin I
5	217.6	61.0	1598	20 AAZ09100	Mistletoe lectin D
6	216.4	60.6	1923	18 AAZ091659	Prepro mistletoe I
7	216.4	60.6	1923	20 AAZ09182	Mistletoe lectin P
8	190.2	53.3	763	20 AAZ09101	Mistletoe lectin A
9	189	52.9	762	20 AAZ09104	Mistletoe lectin A

10	185.8	52.0	756	19 AAV51343	Mistletoe rMLA var
11	185.8	52.0	774	18 AAT91660	Prepro mistletoe 1
12	185.8	52.0	774	20 AAV74180	Mistletoe ML A-cha
13	184.2	51.6	768	22 AAC85474	A-chain gene isofo
14	182.6	51.1	762	22 AAC85472	Mistletoe rMLA DNA
15	182.2	51.0	762	22 AAV51341	A-chain gene isofo
16	182.2	51.0	762	22 AAC85473	Mistletoe lectin A
17	179.4	50.3	768	20 AAZ09105	Mistletoe lectin A
18	90	25.2	1596	20 AAZ09106	Mistletoe lectin I
19	78.6	22.0	762	20 AAZ09107	Mistletoe lectin A
20	78.6	22.0	768	20 AAZ09108	Mistletoe lectin A
21	57	16.0	839	10 AAN91282	Sequence derived f
22	55.4	15.5	1846	8 AAN70526	Sequence of cDNA 1
23	55.2	15.5	1825	22 AAT79921	Mutant preporiclin
24	54.4	15.2	1855	20 AAZ04229	PAP-256 insert con
25	54.4	15.2	1912	20 AAZ04242	PAP-284 insert DNA
26	54.2	15.2	1837	22 AAT77622	Mutant preporiclin
27	54.2	15.2	1843	22 AAT79937	Mutant preporiclin
28	54.2	15.2	1849	22 AAT79933	Mutant preporiclin
29	54	15.1	1855	20 AAZ04233	PAP-266 insert DNA
30	53.8	15.1	807	14 AAQ46085	Sequence encoding
31	53.8	15.1	839	6 AAN50381	Estimated ricin ag
32	53.8	15.1	1050	6 AAN50380	Ricin A sequence 1
33	53.8	15.1	1050	8 AAN70519	Ricin A gene. Esc
34	53.8	15.1	1050	8 AAN70152	Cloning insert PRA
35	53.8	15.1	1050	11 AAN91281	Ricin A encoding 1
36	53.8	15.1	1050	11 AA005786	Partial sequence o
37	53.8	15.1	1140	13 AA027876	Sequence encoding
38	53.8	15.1	1140	14 AAQ46086	Preporiclin gene.
39	53.8	15.1	1605	14 AAN60193	Castor bean prepro
40	53.8	15.1	1695	22 AAT64137	Castor bean prepro
41	53.8	15.1	1698	22 AAT64138	Sequence of cDNA 1
42	53.8	15.1	1731	22 AAT64138	Sequence of cDNA 1
43	53.8	15.1	1801	8 AAN70524	Mutant preporiclin
44	53.8	15.1	1831	22 AAT79905	Mutant preporiclin
45	53.8	15.1	1834	22 AAT79876	Mutant preporiclin

ALIGNMENTS

RESULT 1	
AAC85478	standard; cDNA; 357 BP.
ID	
XX	AAC85478:
AC	
XX	16-MAY-2001 (first entry)
DT	
XX	
DE	Korean mistletoe lectin gene #1.
XX	
XX	Isoform: A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
KW	KML; tumour; KM-110; KML-C; KMHP; KML-ITU; KML-IL;
KW	heparin binding protein; ds.
XX	
OS	Viscum album coloratum.
XX	
PH	Key
FT	Location/Qualifiers
FT	55..57
FT	/*lag= a
FT	/note= "Encodes Ser"
XX	
PN	EP1074560-A2.
XX	
PD	07-FEB-2001.
XX	
XX	27-JUL-2000; 2000EP-0402168.
PF	
XX	27-JUL-1999; 99KR-0030638.
PR	
XX	(MST-) MISTLE BIOTEC CO LTD.
PA	
XX	
PI	Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;

PI Kang T, Park C;
XX
DR WPI: 2001-171044/18.
DR P-PSDB: AAB47096.

PT Novel lectin proteins isolated from Korean mistletoe, useful for
XX enhancing immunity and effectuating anti-tumoral activity -
PS
XX

Claim 30; Page 33; 62pp; English.

CC The sequences given in AAC85478-79 encode lectins isolated from Korean
CC mistletoe. Korean mistletoe lectins (KML) are useful for enhancing
CC immunity and for treating tumours. The KML's are isolated from a
CC protein fraction derived from the leaves, stems and fruits of Korean
CC mistletoe, which is designated KM-110. One of the isolates, KML-C was
CC shown to be effective against colon 26-M3.1 carcinoma and L5178Y-M25
CC lymphoma.

SQ Sequence 357 BP; 77 A; 103 C; 87 G; 86 T; 4 other;

Query Match 98.9%; Score 353; DB 22; Length 357;
Best Local Similarity 100.0%; Pred. No. 6.7e-110;

Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCTNTGTGAGGCTTCGCCGCAATTAAACAGTGGAGTCNTCT 60
Db 1 GCCAGATTCAATCCCATCTNTGTGAGGCTTCGCCGCAATTAAACAGTGGAGTCNTCT 60
QY 61 CCACCAACAATGATACATGCTCGAGCTGAGACAGAGTTGGGGTTCGACATCCACCAAGTC 120
Db 61 CCACCAACAATGATACATGCTCGAGCTGAGACAGAGTTGGGGTTCGACATCCACCAAGTC 120
QY 121 CACAGTCCAGAGTGGCATTTTATACCAATAAGTTGAGATTTCGCCCGTAAC 180
Db 121 CACAGTCCAGAGTGGCATTTTATACCAATAAGTTGAGATTTCGCCCGTAAC 180
QY 181 TTGTGACGTTGACGATGTCGACAGTGTATCTCAGCTTGGCATGATGTTGTTGAA 240
Db 181 TTGTGACGTTGACGATGTCGACAGTGTATCTCAGCTTGGCATGATGTTGTTGAA 240
QY 241 TGCAGTGTGCGGCATTCCTCTCCAGACCACTTCGCCGCTGCTTAAGTCCGTC 300
Db 241 TGCAGTGTGCGGCATTCCTCTCCAGACCACTTCGCCGCTGCTTAAGTCCGTC 300
QY 301 GTGGATGCGGCCAACGATGTCACTGCACTTTTCCGAACCCACCGTGCATGTA 357
Db 301 GTGGATGCGGCCAACGATGTCACTGCACTTTTCCGAACCCACCGTGCATGTA 357

RESULT 2

AAC85479
ID AAC85479 standard; cDNA: 522 BP.

XX AAC85479;

DT 16-MAY-2001 (first entry)

XX Korean mistletoe lectin gene #2.

DE Isoform: A-chain; B-chain; Biosynthesis; lectin; Korean mistletoe;
KW KML; tumour; KM-110; KML-C; KMHP; KML-IIU; KML-III;
KW heparin binding protein; ds.
XX

OS Viscum album coloratum.

PN EPI074560-A2.

PD 07-FEB-2001.

XX 27-JUL-2000; 2000BP-0402168.

XX 27-JUL-1999; 99KR-0030638.

PA (MIST-) MISTLE BIOTECH CO LTD.

XX Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
PI Kang T, Park C;

XX WPI: 2001-171044/18.

DR P-PSDB: AAB47097.

PT Novel lectin proteins isolated from Korean mistletoe, useful for
XX enhancing immunity and effectuating anti-tumoral activity -
PS
XX

Claim 32; Page 34; 62pp; English.

CC The sequences given in AAC85478-79 encode lectins isolated from Korean
CC mistletoe. Korean mistletoe lectins (KML) are useful for enhancing
CC immunity and for treating tumours. The KML's are isolated from a
CC protein fraction derived from the leaves, stems and fruits of Korean
CC mistletoe, which is designated KM-110. One of the isolates, KML-C was
CC shown to be effective against colon 26-M3.1 carcinoma and L5178Y-M25
CC lymphoma.

SQ Sequence 522 BP; 112 A; 150 C; 131 G; 129 T; 0 other;

Query Match 64.4%; Score 230; DB 22; Length 522;
Best Local Similarity 78.7%; Pred. No. 6.1e-68;

Matches 288; Conservative 0; Mismatches 69; Indels 9; Gaps 1;

QY 1 GCCAGATTCAATCCCATCTNTGTGAGGCTTCGCCGCAATTAAACAGTGGAGTCNTCT 60
Db 157 GCCAGATTCAATCCCATCTNTGTGAGGCTTCGCCGCAATTAAACAGTGGAGTCNTCT 216
QY 61 CCACCAACAATGATACATGCTCGAGCTGAGACAGAGTTGGGGTTCGACATCCACCAAGTC 120
Db 217 CTTCCGCAATGATACATGCTCGAGCTGAGACAGAGTTGGGGTTCGACATCCACCAAGTC 276
QY 121 CACAGTCCAGAGTGGCATTTTATACCAATAAGTTGAGATTTCGCCCGTAAC 180
Db 277 CACAGTCCAGAGTGGCATTTTATACCAATAAGTTGAGATTTCGCCCGTAAC 276
QY 181 TTGTGACGTTGACGATGTCGACAGTGTATCTCAGCTTGGCATGATGTTGTTGAA 240
Db 337 TTGTGACGTTGACGATGTCGACAGTGTATCTCAGCTTGGCATGATGTTGTTGAA 396
QY 241 TGCAGTGTGCGGCATTCCTCTCCAGACCACTTCGCCGCTGCTTAAGTCCGTC 300
Db 397 TGTAGGAGACGACCATCTTCTCCAGCTGCGCATTTGGCGCGTGCATACGACCCGTC 456
QY 301 GTGGA-----TGCGCCAACAGATGCACCTGCACTTTTCCGAACCCACCGTGC 351
Db 457 TTGGAAATAGCGCGCGCTGACGATGTACCTGCACTCTTCCGAACCCACCGTGC 516
QY 352 ATCGTA 357
Db 517 ATCGTA 522

RESULT 3

ABL56947
ID ABL56947 standard; DNA: 1656 BP.

XX ABL56947;

DT 08-JUL-2002 (first entry)

XX Galactose-recognising mistletoe lectin encoding polynucleotide.

XX Mistletoe; galactose-recognising mistletoe lectin; MLIII; gene; ds.

OS Viscum album.

XX Key Location/Qualifiers

FT 1..1656
FT CDS
FT /*tag= a

FT /transl_except- (pos:667..669,aa:Val)
FT /transl_except- (pos:751..753,aa:Ile)
FT /transl_except- (pos:1030..1032,aa:Trp)
FT /transl_except- (pos:1138..1140,aa:Val)
FT /transl_except- (pos:1342..1344,aa:Glu)
FT /product- "Galactose-recognising mistletoe lectin"
XX
XX DE10044027-A1.
XX
XX 14-MAR-2002.
XX
XX 06-SEP-2000; 2000DE-1044027.
XX
XX 06-SEP-2000; 2000DE-1044027.
XX
XX (VISC-) VISCUM AG.
XX
XX Kieff S;
XX
XX WPI; 2002-316737/36.
XX
XX P-PSDB; ABB79450.
XX
XX New nucleic acid encoding preprotein of mistletoe lectin, useful as
XX diagnostic and therapeutic agents, also encodes polypeptide -
XX
XX Claim 1: Fig 1; 6pp; German.
XX
XX The invention relates to a nucleic acid molecule (ABU56947) that encodes
XX a preprotein (ABB79450) which, after maturation, has the biological
XX activity of the galactose-recognising mistletoe lectin (MLIII). The MLIII
XX encoding nucleic acid molecule, primers specific to it or complements of
XX it, and encoded (oligomeric) polypeptides are useful as diagnostic and
XX therapeutic agents.
XX
XX Sequence 1656 BP; 387 A; 418 C; 449 G; 402 T; 0 other;
SQ
Query Match 63.7%; Score 227.4; DB 24; Length 1656;
Best Local Similarity 78.4%; Pred. No. 8e-67;
Matches 286; Conservative 0; Mismatches 70; Indels 9; Gaps 1;
XX
XX 1 GCCAGATTCATCCATCMTCTGTGAGGCTTCGCGCAATTAACAGTGGGAGTCTCT 60
DB 598 GCCAGATTCATCCATCMTCTGTGAGGCTTCGCGCAATTAACAGTGGGAGTCTCT 657
XX
XX 61 CCACCAAAATGATGATGCTGAGCTGAGAGAGTGGGGTGCAGATCCACCAAGTGC 120
DB 658 CTTCCAGACATGTACATGCTGAGCTGAGAGAGTGGGGTGCAGATCCACCAAGTGC 717
XX
XX 121 CACGAGTCCAAAGATGATGCTGAGCTGAGAGAGTGGGGTGCAGATCCACCAAGTGC 180
DB 718 CACGAGTCCAAAGATGATGCTGAGCTGAGAGAGTGGGGTGCAGATCCACCAAGTGC 777
XX
XX 181 TTTGTGAGTGTGAGCAATGTTCCGACGTGATCTCCAGTTCGCGATATGTTTCCAA 240
DB 778 TTTGTGAGTGTGAGCAATGTTCCGACGTGATCTCCAGTTCGCGATATGTTTCCAA 837
XX
XX 241 TGCAGTGTGAGCAATGTTCCGACGTGATCTCCAGTTCGCGATATGTTTCCAA 300
DB 838 TGCAGTGTGAGCAATGTTCCGACGTGATCTCCAGTTCGCGATATGTTTCCAA 897
XX
XX 301 GTGGA-----TGCAGCAACAGATGATCAGTTCGAGTTCGAGCCACCGTCCG 351
DB 898 TTGGAATAATAGCGCGCGCTGAGAGATGATCAGTTCGAGTTCGAGCCACCGTCCG 957
XX
XX 352 ATCGT 356
DB 958 ATGCT 962
XX
XX RESULT 4
XX AA209103
XX AA209103 standard; DNA; 1596 BP.
XX

AC AA209103;
XX 18-OCT-1999 (first entry)
XX
XX Mistletoe lectin I DNA fragment.
XX
XX
XX Mistletoe; lectin; antitumor; immunostimulant; A-chain; MLA; immunity;
XX ribozyme 285 subunit; non-cytotoxic; T-cell activation; immune response;
XX lymphokine-producing macrophage; uncontrolled cell growth; treatment;
XX cancer; cytotoxicity; antigen; isoforn; lectin I; ds.
XX
XX Viscum album.
XX
XX DE19804210-A1.
XX
XX 12-AUG-1999.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX 03-FEB-1998; 98DE-1004210.
XX
XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.
XX
XX Morris P, Stiefel T, Voelter W, Welters P;
XX WPI; 1999-445335/38.
XX P-PSDB; AAY25979.
XX
XX Preparation of mistletoe lectins in heterologous systems,
XX particularly for use as anticancer agents and immunostimulants
XX
XX Claim 14: Fig 1A; 78pp; German.
XX
XX This invention describes a novel mistletoe lectin (I) and its fragments
XX which have antitumor and immunostimulatory activity. The A-chain (MLA)
XX of the mistletoe lectin binds to, and inactivates, the 28S subunit of
XX CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
XX CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
XX CC fragments are used to treat uncontrolled cell growth (particularly of
XX CC cancers) and if they lack cytotoxicity, to increase the strength of the
XX CC immune response, particularly to a co-administered antigen
XX (tumour-associated, bacterial or viral). The method allows production of
XX mistletoe lectin, and its individual chains, in many different isoforms
XX and on a large scale, at any time of the year. Recombinant products are
XX CC free from toxins present in natural mistletoe extracts. This sequence
XX encodes a fragment of the mistletoe lectin I protein.
XX
XX Sequence 1596 BP; 387 A; 418 C; 419 G; 372 T; 0 other;
SQ
Query Match 61.1%; Score 218; DB 20; Length 1596;
Best Local Similarity 77.5%; Pred. No. 1.2e-63;
Matches 276; Conservative 0; Mismatches 74; Indels 6; Gaps 1;
XX
XX 1 GCCAGATTCATCCATCMTCTGTGAGGCTTCGCGCAATTAACAGTGGGAGTCTCT 60
DB 499 GCCAGATTCATCCATCMTCTGTGAGGCTTCGCGCAATTAACAGTGGGAGTCTCT 558
XX
XX 61 CCACCAAAATGATGATGCTGAGCTGAGAGAGTGGGGTGCAGATCCACCAAGTGC 120
DB 559 CTTCCAGACATGTACATGCTGAGCTGAGAGAGTGGGGTGCAGATCCACCAAGTGC 618
XX
XX 121 CAGAGTCCAAAGATGATGCTGAGCTGAGAGAGTGGGGTGCAGATCCACCAAGTGC 180
DB 619 CAGAGTCCAAAGATGATGCTGAGCTGAGAGAGTGGGGTGCAGATCCACCAAGTGC 678
XX
XX 181 TTTGTGAGTGTGAGCAATGTTCCGACGTGATCTCCAGTTCGCGATATGTTTCCAA 240
DB 679 TTTGTGAGTGTGAGCAATGTTCCGACGTGATCTCCAGTTCGCGATATGTTTCCAA 738
XX
XX 241 TGCAGTGTGAGCAATGTTCCGACGTGATCTCCAGTTCGCGATATGTTTCCAA 300
DB 739 TGCAGTGTGAGCAATGTTCCGACGTGATCTCCAGTTCGCGATATGTTTCCAA 798
XX

QY 301 GTGATGCGGCAACGATGTCACCTGACACTTTTCCGAACCCAGCGTCGATCGT 356
 Db 799 AT-----AGCCGATGATGTTACTGTCAGTGTCTTGGACACTACGATGCGGATTGT 848

RESULT 5

AAZ09100
 ID AAZ09100 standard; DNA; 1598 BP.

XX
 AC AAZ09100;

DT 18-OCT-1999 (first entry)

DE Mistletoe lectin DNA consensus sequence.

XX Mistletoe; lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; ss.

OS Viscum album.

PN DE19804210-A1.

PD 12-AUG-1999.

PF 03-FEB-1998; 98DE-1004210.

PR 03-FEB-1998; 98DE-1004210.

PA (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

PI Morris P, Stiefel T, Voelter W, Walters P;

DR WPI; 1999-445335/38.

PT Preparation of mistletoe lectins in heterologous systems.

PS Particularly for use as anticancer agents and immunostimulants

XX Claim 11; Page 32-33; 78pp; German.

CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and (II) and its
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of mistletoe lectin DNA described in the
 CC specification.

SO Sequence 1598 BP; 373 A; 396 C; 392 G; 356 T; 81 other;

Query Match 61.0%; Score 217.6; DB 20; Length 1598;
 Best Local Similarity 74.7%; Pred. No. 1,7e-63;

Matches 266; Conservative 13; Mismatches 71; Indels 6; Gaps 1;

QY 1 GCCAATTCATCCCATNTGTGAGGCTTCCGCGCAATTAACAGTGGGAGTCNTCT 60
 Db 500 GCCAATTCATCCCATTTTGTGAGGCKMYCGCAARAVYTTAACAGTGGGAGTCATT 559

QY 61 CCACCAACATGATGATCGATGAGAGCGAGTGGGGTGACACATCCACCAAGTC 120
 Db 560 CTGGCAGACGCTGTACATGCTGAGCTGAGAGAGTGGGCGCAACATCCACGCAAGTC 619

QY 121 CAGCAGTCCAGAGATGTCATTTTAAATACCAATAGATTGAGATTTCCGCGGTATC 180
 Db 620 CAGCATTCAACGATGCGGTTTTTAAATACCAACATYCGTGGCTATATCTACCTGATAC 679

QY 181 TTGTGAGGNTGAGCATGTTGCGAGCTGATGTCAGCTTGGGATCATGTTGCGAA 240
 Db 680 TTGCTGAGCTGTCWCAATAGTTGCGCKMYGTGATGCCACGCTTGGGATCATGTTGTTGTA 739
 QY 241 TGCAGTGGTGGCCATTTCTCTCTCTGACACCCCTTGCGCGGTGCTCTTAAGTCCGTC 300
 Db 740 TCGGAGAGCGGCCCATCTTCTCTGACGTGCGCTATTGCGCGGTGATACGACCCGTG 799

QY 301 GTGATGCGGCAACGATGTCACCTGACACTTTTCCGAACCCAGCGTCGATCGT 356
 Db 800 AT-----AGCCGATGATGTTACTGTCAGTGTCTTGGACACTACGATGCGGATTGT 849

RESULT 6

AAAT91659
 ID AAAT91659 standard; DNA; 1923 BP.

AC AAAT91659;

DT 18-DEC-1997 (first entry)

DE Prepro mistletoe lectin.

KW Mistletoe; lectin; cytotoxic; A chain; B chain; dimer; ss.

OS Viscum album.

PN Key Location/Qualifiers

FT CDS 55..1749

FT misc_feature 1..204

FT misc_feature /*tag= a

FT misc_feature /*tag= b

FT misc_feature /*tag= c

FT misc_feature /*tag= "ML gene fragment f"

FT misc_feature /*tag= "ML gene fragment e"

FT misc_feature /*tag= "ML gene fragment g"

FT misc_feature /*tag= f

FT misc_feature /*tag= "ML gene fragment i"

PN EP751221-A1.

PD 02-JAN-1997.

PF 26-JUN-1995; 95EP-0109949.

PR 26-JUN-1995; 95EP-0109949.

PA (MADU) MADDAUS KOELN AG.

PI Baur A, Eck J, Lentzen H, Zinke H;

DR WPI; 1997-054678/06.

DR P-PSDB; 10021.

DT Nucleic acid encoding pre:pro form of mistletoe lectin - for

PT therapeutic or diagnostic use

PS Claim 1; Fig 4C; 30pp; German.

CC Mistletoe lectin is a cytotoxic agent that has been used for tumour

CC therapy. It can be used in immunotoxins and medicaments. Nucleic

CC acid fragments can be used in diagnostic methods. Mistletoe lectin (

CC AAAT91659) comprises an A chain (AAAT91660) and a B chain (AAAT91661).

XX Sequence 1923 BP; 465 A; 488 C; 495 G; 475 T; 0 other;

PT Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 XX
 XX
 PS Claim 12; Page 34; 78pp; German.

CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC represents a consensus sequence of mistletoe lectin A-chain (MLA) DNA
 CC described in the specification.

XX Sequence 763 BP; 163 A; 215 C; 175 G; 170 T; 40 other;

Query Match 53.3%; Score 190.2; DB 20; Length 763;
 Best Local Similarity 78.4%; Pred. No. 2.6e-54;
 Matches 207; Conservative 13; Mismatches 44; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATCTGTCGAGGCTTCGCCGCAATTACAGTGGGAGTCTCT 60
 DB 500 GCCAGATTCAATCCATCTGTCGAGGCTTCGCCGCAATTACAGTGGGAGTCTCT 60
 QY 61 CCACCAAGATGTACATGCTGCTGAGAGAGAGTGGGCGCATCCACCAAGT 120
 DB 560 CTGCCAGACRTGTACATGCTGAGAGAGAGTGGGCGCATCCACCAAGT 120
 QY 121 CAGCATTCCAGAGATGGCATTTTAAATACCAATTAAGATTGACATTCGCGGTAA 180
 DB 620 CAGCATTCCAGAGATGGCATTTTAAATACCAATTAAGATTGACATTCGCGGTAA 180
 QY 181 TTTGTGACGNTGAGCAATGTCGACGATTCGACATTCGCGGTAA 240
 DB 680 TTTGTGACGNTGAGCAATGTCGACGATTCGACATTCGCGGTAA 240
 QY 241 TGCAGTGGTGGCCATTCCTCT 264
 DB 740 TGCAGTGGTGGCCATTCCTCT 264

RESULT 9
 AAZ09104
 ID AAZ09104 standard; DNA: 762 BP.
 AC AAZ09104;
 XX
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Mistletoe lectin A1 DNA fragment.

XX Mistletoe lectin; antitumour; immunostimulant; A-chain; MLA; immunity;
 KW ribozyme 28S subunit; non-cytotoxic; T-cell activation; immune response;
 KW lymphokine-producing macrophage; uncontrolled cell growth; treatment;
 KW cancer; cytotoxicity; antigen; isoform; lectin A1; ds.
 XX
 OS Viscum album.

XX DE19804210-A1.
 XX
 PD 12-AUG-1999.
 XX
 XX 03-FEB-1998; 98DE-1004210.
 PF
 XX 03-FEB-1998; 98DE-1004210.
 PR
 XX (BIOS-) BIOSYN AR2NEIMITTEL GMBH.

XX
 PI Morris P, Stiefel T, Voelter W, Walters P;
 DR WPI: 1999-445335/38.
 XX
 DR P-PSDB: AAV25980.
 XX
 PT Preparation of mistletoe lectins in heterologous systems,
 PT particularly for use as anticancer agents and immunostimulants
 PS Claim 15; Fig 2A; 78pp; German.

CC This invention describes a novel mistletoe lectin (I) and its fragments
 CC which have antitumour and immunostimulatory activity. The A-chain (MLA)
 CC of the mistletoe lectin binds to, and inactivates, the 28S subunit of
 CC ribosomes. Non-cytotoxic forms of (I) activate T-cell and
 CC lymphokine-producing macrophages, so stimulate immunity. (I) and its
 CC fragments are used to treat uncontrolled cell growth (particularly
 CC cancers) and if they lack cytotoxicity, to increase the strength of the
 CC immune response, particularly to a co-administered antigen
 CC (tumour-associated, bacterial or viral). The method allows production of
 CC mistletoe lectin, and its individual chains, in many different isoforms
 CC and on a large scale, at any time of the year. Recombinant products are
 CC free from toxins present in natural mistletoe extracts. This sequence
 CC encodes a fragment of the mistletoe lectin A1 protein.

XX Sequence 762 BP; 171 A; 230 C; 183 G; 178 T; 0 other;

Query Match 52.9%; Score 189; DB 20; Length 762;
 Best Local Similarity 81.8%; Pred. No. 6.6e-54;
 Matches 216; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 GCCAGTTTAAATCCATCTGTCGAGGCTTCGCCGCAATTACAGTGGGAGTCTCT 60
 DB 499 GCCAGTTTAAATCCATCTGTCGAGGCTTCGCCGCAATTACAGTGGGAGTCTCT 60
 QY 61 CCACCAAGATGTACATGCTGCTGAGAGAGAGTGGGCGCATCCACCAAGT 120
 DB 559 CTGCCAGACRTGTACATGCTGAGAGAGAGTGGGCGCATCCACCAAGT 120
 QY 121 CAGCATTCCAGAGATGGCATTTTAAATACCAATTAAGATTGACATTCGCGGTAA 180
 DB 619 CAGCATTCCAGAGATGGCATTTTAAATACCAATTAAGATTGACATTCGCGGTAA 180
 QY 181 TTTGTGACGNTGAGCAATGTCGACGATTCGACATTCGCGGTAA 240
 DB 679 TTTGTGACGNTGAGCAATGTCGACGATTCGACATTCGCGGTAA 240
 QY 241 TGCAGTGGTGGCCATTCCTCT 264
 DB 739 TGCAGTGGTGGCCATTCCTCT 264

RESULT 10
 AAV51343
 ID AAV51343 standard; DNA: 756 BP.
 AC AAV51343;
 XX
 XX
 DT 23-OCT-1998 (first entry)
 XX
 DE Mistletoe rMLA variant DNA.

XX Lectin A-chain; mistletoe; rMLA; fusion protein; effector; cytotoxic;
 KW intracellular; processing module; protease recognition; targeting module;
 KW internalisation; treatment; disorder; cell proliferation; activation;
 KW autoimmune disease; allergy; tumour; ricin; translocation; ss.
 XX
 OS Viscum album.

XX
 FH Key Location/Qualifiers
 FT 1..756
 FT CDS
 FT /product="Lectin A-chain"

FT /note="Partial sequence"
 XX MO9829540-A2.
 PN 09-JUL-1998.
 PD 02-JAN-1998; 98WO-EP00009.
 PF 02-JAN-1997; 97EP-0100012.
 PR 02-JAN-1997; 97EP-0100012.
 XX (BRAT-) BRAIN BIOTECHNOLOGY RES & INFORMATION NE.
 PA Eck J, Schmidt A, Zinke H;
 PI WPI: 1998-388122/33.
 DR Nucleic acid encoding fusion protein containing mistletoe lectin A
 XX chain - useful for treatment of proliferative and autoimmune
 PT diseases, allergies and tumours
 PS Disclosure; Fig 11a'; 115pp; German.

XX This sequence encodes a variant mistletoe lectin A-chain, rMLA. This
 CC sequence can be used in the construction of a fusion protein which
 CC comprises an effector module that is cytotoxic intracellularly, a
 CC processing module covalently bonded to the effector module and
 CC containing a protease recognition sequence, and a targeting module
 CC covalently bonded to the processing module, able to bind specifically to
 CC the surface of a cell so as to mediate internalisation of the fusion
 CC protein. Such a fusion protein can be used for treating disorders
 CC involving proliferation and/or elevated activation of cells, especially
 CC autoimmune disease, allergy and tumours. The proteins can be administered
 CC e.g. by injection or topically but especially by intravenous injection,
 CC at 1 ng to 500 mu g/kg/day, or for ex vivo use at 1 pg to 500 ng/ml.
 CC Fusion proteins can develop toxic activity in a wide range of target
 CC cells. The processing module prevents extracellular dissociation, and
 CC fusion proteins based on mistletoe lectin A-chain are far more active
 CC than those based on ricin and do have the associated problems of
 CC non-specific toxicity. The protein may be expressed in a non-glycosylated
 CC form that does not bind to sugar receptors in the liver, and which has a
 CC long half-life in the blood. Where the mistletoe lectin B-chain is used,
 CC it actively assists in translocation of the ML A-chain from the
 CC endoplasmic reticulum to the cytoplasm.
 CC
 XX Sequence 756 BP; 167 A; 229 C; 181 G; 179 T; 0 other;
 SQ

Query Match 52.0%; Score 185.8; DB 19; Length 756;
 Best Local Similarity 82.4%; Pred. No. 8.1e-53;
 Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

OY 1 GCCAGATTCAATCCCATCTGTCGAGGCTTCCCGCAATTAACAGTGGGAGTCTCT 60
 DB 499 GCCAGATTCAATCCCATCTGTCGAGGCTTCCCGCAATTAACAGTGGGAGTCTCT 558
 OY 61 CCACAAACATGTACATGCTGAGCTGAGAGAGTGGGGTGCACATCCACCAAGTTC 120
 DB 559 CTGCCAGACTGTACATGCTGAGCTGAGAGAGTGGGGTGCACATCCACCAAGTTC 618
 OY 121 CAGCAGTCCAGAGTGGCATTTTATACCAATTAAGTTGCAGATTTCCCGCGGTAC 180
 DB 619 CAGATTCAACAGTGGCGGCTTTTATTAACCAATTCGTTGCTTACCCCGCGGTAC 678
 OY 181 TTTGTGACGNTGAGCAATGTTCCGACGTGATCTCCAGTTGGCGATCATGTTGTGCA 240
 DB 679 TTGTTGACGTTGACCAATGTTCCGACGTGATCTCCAGTTGGCGATCATGTTGTGCA 738
 OY 241 TGCAGTGGTGGCCAT 256
 DB 739 TGCAGAGAGCGGCCAT 754

RESULT 11
 AAT91660

ID AAT91660 standard; DNA; 774 BP.

AC AAT91660;

DT 18-DEC-1997 (first entry)

DE Prepro mistletoe lectin A chain.

KW Mistletoe; lectin; cytotoxic; A chain; B chain; dimer; ss.

XX Viscum album.

FT Key Location/Qualifiers

FT CDS 4..765

FT /tag= a

PN EP751221-A1.

PD 02-JAN-1997.

PP 26-JUN-1995; 95EP-0109949.

PR 26-JUN-1995; 95EP-0109949.

PA (MADU) MADAVUS KOELN AG.

PI Baur A, Eck J, Lentzen H, Zinke H;

DR WPI: 1997-054678/06.

DR P-PSDB; AAM10022.

PT Nucleic acid encoding pre-pro form of mistletoe lectin - for
 PT therapeutic or diagnostic use

PS Claim 2; Fig 4A; 30pp; German.

XX Mistletoe lectin is a cytotoxic agent that has been used for tumour

CC therapy. It can be used in immunotoxic and medicaments. Nucleic

CC acid fragments can be used in diagnostic methods. Mistletoe lectin (

CC AAT91659) comprises an A chain (AAT91660) and a B chain (AAT91661).

XX Sequence 774 BP; 173 A; 232 C; 185 G; 184 T; 0 other;
 SQ

Query Match 52.0%; Score 185.8; DB 18; Length 774;
 Best Local Similarity 82.4%; Pred. No. 8.2e-53;
 Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

OY 1 GCCAGATTCAATCCCATCTGTCGAGGCTTCCCGCAATTAACAGTGGGAGTCTCT 60
 DB 505 GCCAGATTCAATCCCATCTGTCGAGGCTTCCCGCAATTAACAGTGGGAGTCTCT 564
 OY 61 CCACAAACATGTACATGCTGAGCTGAGAGAGTGGGGTGCACATCCACCAAGTTC 120
 DB 565 CTGCCAGACTGTACATGCTGAGCTGAGAGAGTGGGGTGCACATCCACCAAGTTC 624
 OY 121 CAGCAGTCCAGAGTGGCATTTTATACCAATTAAGTTGCAGATTTCCCGCGGTAC 180
 DB 625 CAGATTCAACAGTGGCGGCTTTTATTAACCAATTCGTTGCTTACCCCGCGGTAC 684
 OY 181 TTTGTGACGNTGAGCAATGTTCCGACGTGATCTCCAGTTGGCGATCATGTTGTGCA 240
 DB 685 TTGTTGACGTTGACCAATGTTCCGACGTGATCTCCAGTTGGCGATCATGTTGTGCA 744
 OY 241 TGCAGTGGTGGCCAT 256
 DB 745 TGCAGAGAGCGGCCAT 760

RESULT 12
 ID AAT91660 standard; DNA; 774 BP.
 ID AAT91660
 AC AAT91660;

XX 30-APR-1999 (first entry)
 DT Mistletoe ML A-chain DNA.
 XX
 DE
 XX
 KW ML: mistletoe; lectin; MLA; A-chain; transgenic plant; glycosylation;
 KW dimer; immunotoxin; large-scale production; diagnosis; therapeutic;
 KW cancer; ss.
 XX
 OS Viscum album.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 4..765
 FT /*tag= a
 FT /product= "mistletoe lectin"
 XX
 PN EP884388-A1.
 XX
 PD 16-DEC-1998.
 XX
 PE 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 95EP-0109949.
 XX
 PR 26-JUN-1995; 98EP-0105660.
 XX
 PA (MADU) MADAU KOELN AG.
 XX
 PI Baur A, Eck J, Lentzen H, Zinke H;
 XX
 DR WPI; 1999-026582/03.
 DR P-PSDB; AAW90125.
 XX
 PT New transgenic plant expressing mistletoe lectin - useful for
 PT producing recombinant lectin in e.g. cancer diagnosis and therapy
 XX
 PS Claim 2; Fig 4a; 30pp; German.
 XX
 CC This invention describes a novel transgenic plant transformed with a
 CC vector capable of encoding a mistletoe (Viscum album) lectin
 CC preproprotein or a biologically active fragment. The specification
 CC also describes a polypeptide produced by a plant where the polypeptide
 CC exhibits at least one enzymatic modification other than the glycosylation
 CC that occurs in viscum album or the polypeptide is a fusion protein, a
 CC mistletoe lectin polypeptide dimer and an immunotoxin comprising the
 CC polypeptide or the polypeptide dimer. The plants are used for large-scale
 CC production of mistletoe lectin for diagnostic or therapeutic purposes
 CC (e.g. in cancer therapy). This sequence encodes the mistletoe lectin
 CC A-chain which is contained in expression vector pT7MA.
 CC
 XX
 SQ Sequence 774 BP; 173 A; 232 C; 185 G; 184 T; 0 other;
 Query Match 52.0%; Score 185.8; DB 20; Length 774;
 Best Local Similarity 82.4%; Pred. No. 8.2e-53;
 Matches 211; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 1 GCCAATTCAATCCCATNTGTGAGGCTTCGCCGAATTAACAGTGGGAGTCNTCT 60
 DB 505 GCCAATTCAATCCCATNTGTGAGGCTTCGCCGAATTAACAGTGGGAGTCATTTT 564
 QY 61 CCACCAACATGTATCATGCTGAGCTGAGAGCGATGGGGGTGCACAATCCACCAAGTC 120
 DB 565 CTGCGACGCTGTACATGCTGAGCTGAGAGCGATGGGGGTGCACAATCCACCAAGTC 624
 QY 121 CAGCAGTCCAAAGATGGCATTTTAAATACCAATAGATTGCAAGTTTCCCGCGTAAC 180
 DB 625 CAGCAGTCCAAAGATGGCGTTTAAATACCAATTCGGTTGCTATACCCCGCGTAAC 684
 QY 181 TTGTGACGNTGACCAATGTTTCGAGCGATGTCGAGCTTGGCGATCATGTTGTCGAA 240
 DB 685 TTGCTGTGTTGAGCAATGTTTCGAGCGATGTCGAGCTTGGCGATCATGTTGTCGAA 744
 QY 241 TGCAGTGTGGGCGCAT 256
 DB 241 TGCAGTGTGGGCGCAT 256

DB 745 TCGGAGAGCGGCCAT 760
 RESULT 13
 AAC85474
 ID AAC85474 standard; cDNA; 768 BP.
 XX
 AC AAC85474;
 XX
 AC
 XX
 DT 16-MAY-2001 (first entry)
 DE
 XX
 DE A-chain gene isoform for biosynthesis of a Korean mistletoe lectin #3.
 KW Isoform; A-chain; B-chain; biosynthesis; lectin; Korean mistletoe;
 KW KML; tumour; KM-110; KML-C; KMHP; KML-IIU; KML-III;
 KW heparin binding protein; ds.
 XX
 OS Viscum album coloratum.
 XX
 PN EP1074560-A2.
 XX
 PD 07-FEB-2001.
 XX
 PE 27-JUL-2000; 2000EP-0402168.
 XX
 PR 27-JUL-1999; 99KR-0030638.
 XX
 PA (MIST-) MISTLE BIOTECH CO LTD.
 XX
 PI Kim J, Song S, Suh B, Lee K, Doo M, Kwak J, Song B, Yoon T;
 PI Kang T, Park C;
 XX
 DR WPI; 2001-171044/18.
 DR P-PSDB; AAB47092.
 XX
 PT Novel lectin proteins isolated from Korean mistletoe, useful for
 PT enhancing immunity and effectuating anti-tumoral activity -
 XX
 PS Claim 5; Page 27-28; 62pp; English.
 XX
 CC The sequences given in AAC85472-74 encode isoforms of an A-chain gene
 CC involved in biosynthesis of lectins isolated from Korean mistletoe.
 CC Korean mistletoe lectins (KML) are useful for enhancing immunity and
 CC for treating tumors. The KML's are isolated from a protein fraction
 CC derived from the leaves, stems and fruits of Korean mistletoe, which
 CC is designated KM-110. One of the isolates, KML-C was shown to be
 CC effective against colon 26-M3.1 carcinoma and L5178Y-ML25 lymphoma.
 CC
 XX
 SQ Sequence 768 BP; 171 A; 219 C; 200 G; 178 T; 0 other;
 Query Match 51.6%; Score 184.2; DB 22; Length 768;
 Best Local Similarity 80.7%; Pred. No. 2.9e-52;
 Matches 213; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 1 GCCAATTCAATCCCATNTGTGAGGCTTCGCCGAATTAACAGTGGGAGTCNTCT 60
 DB 505 GCCAATTCAATCCCATNTGTGAGGCTTCGCCGAATTAACAGTGGGAGTCATTTT 564
 QY 61 CCACCAACATGTATCATGCTGAGCTGAGAGCGATGGGGGTGCACAATCCACCAAGTC 120
 DB 565 CTGCGACGCTGTACATGCTGAGCTGAGAGCGATGGGGGTGCACAATCCACCAAGTC 624
 QY 121 CAGCAGTCCAAAGATGGCATTTTAAATACCAATAGATTGCAAGTTTCCCGCGTAAC 180
 DB 625 CAGCAGTCCAAAGATGGCGTTTAAATACCAATTCGGTTGCTATACCCCGCGTAAC 684
 QY 181 TTGTGACGNTGACCAATGTTTCGAGCGATGTCGAGCTTGGCGATCATGTTGTCGAA 240
 DB 685 TTGCTGTGTTGAGCAATGTTTCGAGCGATGTCGAGCTTGGCGATCATGTTGTCGAA 744
 QY 241 TGCAGTGTGGGCGCATTCCTCT 264
 DB 745 TGCAGTGTGGGCGCATTCCTCT 768

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2003, 18:18:22 ; Search time 21.5276 Seconds
(without alignments)
5085.723 Million cell updates/sec

Title: US-09-627-165e-13

Sequence: 357
1 gccagatccaatccatcctt.....aacccacccgtcgatcgta 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216.4	60.6	1923	4 US-08-776-059-34	Sequence 34, Appl
2	185.8	52.0	705	4 US-08-776-059-38	Sequence 38, Appl
3	185.8	52.0	774	4 US-08-776-059-30	Sequence 30, Appl
4	53.8	15.1	807	2 US-08-356-786-7	Sequence 7, Appl
5	53.8	15.1	1140	1 US-08-218-303-15	Sequence 15, Appl
6	53.8	15.1	1140	2 US-08-338-793D-60	Sequence 60, Appl
7	53.8	15.1	1605	2 US-08-356-786-9	Sequence 9, Appl
8	53.8	15.1	1855	4 US-09-147-208-33	Sequence 33, Appl
9	53.8	15.1	1855	4 US-09-147-208-47	Sequence 47, Appl
10	53.8	15.1	1855	4 US-09-147-208-54	Sequence 54, Appl
11	53.8	15.1	1855	4 US-09-147-208-40	Sequence 40, Appl
12	53.8	15.1	1879	4 US-09-147-208-23	Sequence 23, Appl
13	53.8	15.1	1879	4 US-09-147-208-24	Sequence 24, Appl
14	53.8	15.1	1879	4 US-09-147-208-25	Sequence 25, Appl
15	31.6	8.9	2607	2 US-08-907-166-1	Sequence 1, Appl
16	29.6	8.3	676	4 US-08-998-416-1130	Sequence 1130, Ap
17	29.6	8.3	724	4 US-08-998-416-928	Sequence 928, Ap
18	29.6	8.3	31728	4 US-09-453-702B-64	Sequence 64, Appl
19	29.2	8.2	3747	2 US-09-080-887-1	Sequence 1, Appl
20	29.2	8.2	3747	4 US-09-323-735-1	Sequence 1, Appl
21	29.2	8.2	5822	3 US-08-899-595-4	Sequence 4, Appl
22	29.2	8.2	5822	3 US-08-899-595-5	Sequence 5, Appl
23	29	8.1	4403765	4 US-09-103-840A-2	Sequence 2, Appl
24	28.8	8.1	1495	4 US-09-063-898-1	Sequence 1, Appl
25	28.6	8.0	1207	1 US-08-460-806-16	Sequence 16, Appl
26	28.6	8.0	1207	1 US-08-325-630-16	Sequence 16, Appl
27	28.6	8.0	1353	4 US-09-518-657-3	Sequence 3, Appl

28	28.6	8.0	1598	4 US-09-518-657-6	Sequence 6, Appl
29	28.4	8.0	9069	4 US-08-961-527-97	Sequence 97, Appl
30	27.8	7.8	2348	4 US-08-990-823-61	Sequence 61, Appl
31	27.8	7.8	36519	3 US-08-923-137-2	Sequence 2, Appl
32	27.6	7.7	391	3 US-08-444-818-80	Sequence 80, Appl
33	27.6	7.7	480	3 US-08-444-818-164	Sequence 164, App
34	27.6	7.7	789	4 US-08-776-059-42	Sequence 42, Appl
35	27.6	7.7	807	4 US-08-776-059-32	Sequence 32, Appl
36	27.6	7.7	1019	3 US-08-444-818-154	Sequence 154, App
37	27.6	7.7	1207	1 US-08-460-806-12	Sequence 12, Appl
38	27.6	7.7	1207	1 US-08-325-630-12	Sequence 12, Appl
39	27.6	7.7	1210	2 US-08-483-695-35	Sequence 35, Appl
40	27.6	7.7	1210	2 US-07-965-285-35	Sequence 35, Appl
41	27.6	7.7	1210	2 US-08-487-231-35	Sequence 35, Appl
42	27.6	7.7	1210	4 US-09-201-912-35	Sequence 35, Appl
43	27.6	7.7	1989	3 US-08-824-057-4	Sequence 4, Appl
44	27.6	7.7	1989	4 US-09-415-582-4	Sequence 4, Appl
45	27.6	7.7	4378	2 US-09-080-887-3	Sequence 3, Appl

ALIGNMENTS

```
RESULT 1
US-08-776-059-34
; Sequence 34, Application US/08776059B
; Patent No. 6271368
; GENERAL INFORMATION:
; APPLICANT: LENTZEN, Hans
; APPLICANT: ECK, Jurgen
; APPLICANT: BAUR, Axel
; APPLICANT: ZINKE, Holger
; TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
; FILE REFERENCE: 674503-2003
; CURRENT APPLICATION NUMBER: US/08/776, 059B
; EARLIER FILING DATE: 1999-06-19
; EARLIER APPLICATION NUMBER: PCT/EP96/02273
; EARLIER FILING DATE: 1996-06-25
; EARLIER APPLICATION NUMBER: 95109949.8
; EARLIER FILING DATE: 1995-06-26
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 34
; LENGTH: 1923
; TYPE: DNA
; ORGANISM: Viscum album
US-08-776-059-34

Query Match      60.6%; Score 216.4; DB 4; Length 1923;
Best Local Similarity 77.2%; Pred. No. 2.9e+66;
Matches 275; Conservative 0; Mismatches 75; Indels 6; Gaps 1.1;

QY 1 GCCAGATTCATCCATCMTGTGAGGCTTCCGCCCAATTAACAGTGGAGTCNTCT 60
   |||||||
DB 652 GCCAGATTCATCCATCMTGTGAGGCTTCCGCCCAATTAACAGTGGAGTCNTCT 711
   |||||||

QY 61 CCAACAATGATCATGCTGAGCTGAGAGAGAGTTGGGCTGCACATCCACCAATC 120
   |||||||
DB 712 CTGCGAACGCTGATGCTGAGCTGAGAGAGAGTTGGGCTGCACATCCACCAATC 771
   |||||||

QY 121 CAGCATTCACAGATGCTGATTTTAACCAATAAATGATTCGAGATTCGCGGTAAC 180
   |||||||
DB 772 CAGCATTCACAGATGCTGATTTTAACCAATAAATGATTCGAGATTCGCGGTAAC 831
   |||||||

QY 181 TTTCGAGGNTGACATGTTTCGAGAGTATCTCAGCTTGGGATCATGTTGCGAA 240
   |||||||
DB 832 TTTCGAGGNTGACATGTTTCGAGAGTATCTCAGCTTGGGATCATGTTGCGAA 891
   |||||||

QY 241 TGCAGTGGTGGCATTCCTCTCTCTGACACACCTTGGCGCTCTTAAGTCCGTC 300
   |||||||
DB 892 TGCAGTGGTGGCATTCCTCTCTCTGAGTGGCGCTCTTAAGTGGCGCTCATACACCGTC 951
   |||||||

QY 301 GTGATGGCGCAACGATGTCNCTGCACTTTTCGACACCAACCGTGGCGATCGT 356
   |||||||
```

Db 952 AT-----AGCGATGATGTTACCTGCAGTCTTCGGAACCTACGGTCCGSAATTGT 1001

RESULT 2

```

US-08-776-059-38
Sequence 38, Application US/08776059B
Patent No. 6271368
GENERAL INFORMATION:
APPLICANT: LENTZEN, Hans
APPLICANT: ECK, Jürgen
APPLICANT: BAUR, Axel
APPLICANT: ZINKE, Holger
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
FILE REFERENCE: 674503-003
CURRENT APPLICATION NUMBER: US/08/776, 059B
CURRENT FILING DATE: 1999-06-19
EARLIER APPLICATION NUMBER: PCT/EP96/02273
EARLIER FILING DATE: 1996-06-25
EARLIER APPLICATION NUMBER: 95109949.8
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 38
LENGTH: 705
TYPE: DNA
ORGANISM: Viscum album
US-08-776-059-38

```

Query Match	52.0%	Score 185.8;	DB 4;	Length 705;
Best Local Similarity	82.48;	Pred. No. 1e-55;		
Matches 211; Conservative	0;	Mismatches 45;	Indels 0;	Caps 0

QY	1	GCNATTCATCCCATCTNTGTGAGAGCTTGCCGGCAAAATTAACAGTGGGAGTCA	60
Db	448	GCCAAATTCATCCCATCTTATTTAGAGGGCTGCCAAATCAATTAACAGTGGGAGTCA	507
QY	61	CCACCAAAACATGTACATGCTGAGCTGGAGAGACAGATTGGGTTCGACAAATCCACCAAGTC	120
Db	508	CTGGCAGACGTGTACATGCTGAGACTGAGACAGATTGGGCCCAACATTCACGCAAGTC	567
QY	121	CAGACGTCACAGATGGGATTTTAAATACCAAAATTAAGATTGCAGATTTCCGCCGGTAAC	180
Db	568	CAGCATTAACCGAATGGGGTTTTTAAATAAACCAATTTGGGTGGCTTAACCCCCGGTAAC	627
QY	181	TTTGTGACGNTGACAAATGTCGGAGGTGATCTCAACGCTGGGGATCATGTTGTGAA	240
Db	628	TTTCGTGACGTGACCAATGTTCCGGACACTGATCCCAAGCTTGGGATCATGTTGTGTA	687
QY	241	TGCAGTGGTGGCCAT	256
Db	688	TGCGGAGAGCGGCCAT	703

RESULT 3

US-08-776-059-30
Sequence 30, Application US/08776059B
Patent No. 6271368
GENERAL INFORMATION:
APPLICANT: LEMTZEN, Hans
APPLICANT: ECK, Jürgen
APPLICANT: BAUR, Axel
APPLICANT: ZINKE, Holger
TITLE OF INVENTION: RECOMBINANT MISTLETOE LECTIN (RML)
FILE REFERENCE: 674503-0003
CURRENT APPLICATION NUMBER: US/08/776,059B
CURRENT FILING DATE: 1999-06-19
EARLIER APPLICATION NUMBER: PCT/EP96/02273
EARLIER FILING DATE: 1996-06-25
EARLIER APPLICATION NUMBER: 95109949.8
EARLIER FILING DATE: 1995-06-26
NUMBER OF SEQ ID NOS: 56
SOFTWARE: Patentln Ver. 2.0

```

; SEQ ID NO 30
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Viscum album
; US-08-776-059-30

```

Query Match	52.0%;	Score 185.8;	DB 4;	Length 774;
Best Local Similarity	82.4%;	Pred. No. 1.1e-55;		
Matches 21;	Conservative 0;	Mismatches 45;	Indels 0;	Gaps 0;

QY	1	GCAGATTCGAATCCCATCTMTGGAGGCTTGCCCGCAATTACAGTGGGAGTCNCT	60
Db	505	GCCAGATTCGAATCCCATCTTATGAGAGGCTTCGCAATACATTACGTGGGCGTCATTT	56
QY	61	CCACCAACAATGATGATGCTGAGCTGAGACGAGATTTGGGGTCGCAATACACCNAATGC	120
Db	565	CTCCCGAGAGCTGATACATCTGTGAGCTGAGACGAGTGTGGGGCAACATTCACGCAATGC	62
QY	121	CACGACATTCGAAGATGGGCAATTTTAATATCCCAATTAAGATTCGACATTTTCGCGCGTAAC	180
Db	625	CACATTTCAACCGATGAGGGCTTTTAATATACCAATTCGGTGTGGCATATACCCCGGTATAC	68
QY	181	TTTGTGACAGTGAAGCAATGTTCCGAGCGATATCCAGCTTGGCGATCATGTTGTGGAA	240
Db	685	TTTGTCGACGTTGACCAATGTTCCGAGCGTATGCCAGCTTGGCGATCATGTTGTGTA	744
QY	241	TTCAGTGGTGGCCAT	256
Db	745	TGGGAGAGCGGCCAT	760

RESULT 4

US-08-356-786-7
Sequence 7, Application US/08356786
Patent No. 5877305
GENERAL INFORMATION:
APPLICANT: Huscon, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. I.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 807 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..807
OTHER INFORMATION: /note="product = "Ricin-A chain"
US-08-356-786-7

Query Match 15.1%; Score 53.8; DB 2; Length 807;
Best Local Similarity 50.0%; Pred. No. 3.5e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCTGTCGAGGCTTGGCCGCAAAATTAACAGTGGGAGTCTCT 60
DB 538 GCAGATTTCATATATTTGAGGAGAAATGCGCAGAGAAATTAAGTACACCGAGATCT 597
QY 61 CCACCAACATGTACATGCTGAGCTGGAGACGAGAGTGGGTCGACAAATCCACCAAGTC 120
DB 598 GCACGATCTAGCTAGCTATTAACACTTGAATAGTGGGAGACTTTCACACTGCATTT 657
QY 121 CAGCACTCCAGAGAGGCAATTTTAAATCCCAATAGATTGACAGATTTCCGCGGTAC 180
DB 658 CAAGAGCTTAACCAAGAGACCTTGTCTAGTCCAAATTCACACTGCAAAAGAGTAATGTTCC 717
QY 181 TTGTGACGNTGAGCAATGTTCGCGACGTGATCTCCAGCTTGGCGATCATGTTGTCGAA 240
DB 718 AATTGAGTGTAGCATGTGAGTATATTAATCCATCATAGCTCATGGTGTATAGA 777
QY 241 TGCAGTGTGCGCCATTCTC 260
DB 778 TGGCAGCTTCACCATCTGTC 797

RESULT 5

US-08-218-303-15
Sequence 15, Application US/08218303
Patent No. 5547867
GENERAL INFORMATION:
APPLICANT: Kara, Bhuphendra V.
APPLICANT: Hockney, Robert C.
APPLICANT: Fitton, John E.
TITLE OF INVENTION: FERMENTATION PROCESS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1615 L Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,303
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/841,533
FILING DATE: 26-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: PNK/3893/94908/PAW.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:

LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 143..943
US-08-218-303-15

Query Match 15.1%; Score 53.8; DB 1; Length 1140;
Best Local Similarity 50.0%; Pred. No. 4.2e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCCATCTGTCGAGGCTTGGCCGCAAAATTAACAGTGGGAGTCTCT 60
DB 677 GCAGATTTCATATATTTGAGGAGAAATGCGCAGAGAAATTAAGTACACCGAGATCT 736
QY 61 CCACCAACATGTACATGCTGAGCTGGAGACGAGAGTGGGTCGACAAATCCACCAAGTC 120
DB 737 GCACGATCTAGCTAGCTATTAACACTTGAATAGTGGGAGACTTTCACACTGCATTT 796
QY 121 CAGCACTCCAGAGAGGCAATTTTAAATCCCAATAGATTGACAGATTTCCGCGGTAC 180
DB 797 CAAGAGCTTAACCAAGAGACCTTGTCTAGTCCAAATTCACACTGCAAAAGAGTAATGTTCC 856
QY 181 TTGTGACGNTGAGCAATGTTCGCGACGTGATCTCCAGCTTGGCGATCATGTTGTCGAA 240
DB 857 AATTGAGTGTAGCATGTGAGTATATTAATCCATCATAGCTCATGGTGTATAGA 916
QY 241 TGCAGTGTGCGCCATTCTC 260
DB 917 TGGCAGCTTCACCATCTGTC 936

RESULT 6

US-08-338-793D-60
Sequence 60, Application US/08338793D
Patent No. 5840521
GENERAL INFORMATION:
APPLICANT: Barth, Peter Thomas
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY CUSHMAN
ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC/XT/AT Compatibles
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word or ASCII editors
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,793D
FILING DATE: 08-NO. 5840521-94
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842,081
FILING DATE: 26-FEB-92
CLASSIFICATION: 435
APPLICATION NUMBER: 9104017.0
FILING DATE: 26-FEB-91
APPLICATION NUMBER: 9109188.4
FILING DATE: 29-APR-91
ATTORNEY/AGENT INFORMATION:
NAME: Kokulis, Paul N.
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: DJB/9901/215431/TGW

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140
TYPE: Nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-338-793D-60

Query Match 15.1%; Score 53.8; DB 2; Length 1140;
Best Local Similarity 50.0%; Pred. No. 4.2e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATGTCGAGAGGCTTCGCCGCAATTAACAGTGGAGTCNTCT 60
Db 677 GCAGAGTCCATATATTTAGGAGGAGAAATGCCAGAGATTAGGTACAAACCGAGATCT 736
QY 61 CCACCAACATGTATGCTCGAGCTGAGAGAGAGTGGGTGACAAATCCACCAAGTC 120
Db 737 GCACCGAGATCCCTAGCTAATTAAGTACCTGAGATAGTGGGAGACTTCCACTGCAAT 796
QY 121 CACGAGTCCAGAGATGGCAATTTTAAATACCAATTAAGTATTCGAGATTTCCGCGGTAC 180
Db 797 CAAGATCTACCAAGAGACCTTTGCTAGTCCAAATTCAGAAAGAGTAAATGGTTCC 856
QY 181 TTTGTAGCAGTGTAGCAATGTTGCGAGCTGATCTCCAGCTTGGCGATGTTGTTGGA 240
Db 857 AAATTCAGTGTAGCATGTAGTATTAATTAATCCATATCATAGCTCTCATGTATAGA 916
QY 241 TGCAGTGTGCGGCATTCCTC 260
Db 917 TCGCAGCTTCACCATCGTC 936

RESULT 7

US-08-356-786-9
Sequence 9, Application US/08356786
Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27, 829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS

LOCATION: 1..1605

OTHER INFORMATION: /note= "product = "G-Fir""

Query Match 15.1%; Score 53.8; DB 2; Length 1605;
Best Local Similarity 50.0%; Pred. No. 5.1e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCAATCCATGTCGAGAGGCTTCGCCGCAATTAACAGTGGAGTCNTCT 60
Db 544 GCAGAGTCCATATATTTAGGAGGAGAAATGCCAGAGATTAGGTACAAACCGAGATCT 603
QY 61 CCACCAACATGTATGCTCGAGCTGAGAGAGAGTGGGTGACAAATCCACCAAGTC 120
Db 604 GCACCGAGATCCCTAGCTAATTAAGTACCTGAGATAGTGGGAGACTTCCACTGCAAT 663
QY 121 CACGAGTCCAGAGATGGCAATTTTAAATACCAATTAAGTATTCGAGATTTCCGCGGTAC 180
Db 664 CAAGATCTACCAAGAGACCTTTGCTAGTCCAAATTCAGAAAGAGTAAATGGTTCC 723
QY 181 TTTGTAGCAGTGTAGCAATGTTGCGAGCTGATCTCCAGCTTGGCGATGTTGTTGGA 240
Db 724 AAATTCAGTGTAGCATGTAGTATTAATTAATCCATATCATAGCTCTCATGTATAGA 783
QY 241 TGCAGTGTGCGGCATTCCTC 260
Db 784 TCGCAGCTTCACCATCGTC 803

RESULT 8

US-09-147-208-33
Sequence 33, Application US/09147208
Patent No. 6333303

GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: Antiviral Ricin-Like Proteins
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/147,208
FILING DATE: 02-MAR-1999
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Rudolph, John R.
REGISTRATION NUMBER: 38, 003
REFERENCE/DOCKET NUMBER: 7841-76

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:

61 CCACCAACATGTACATGCTCGAGCTGAGAGACGAGTTGGGGTGGACAATCCACCAAGTC 120

Db 706 GCACCGAGATCCTAGCCTAATTACCTTGAGATAGTGGGAGACTTCCACTCAATT 765
QY 121 CAGCAGTCCAAAGATGATTTTATATACCAAAATTAAGATTCCAGATTCGCGGTAC 180
Db 766 CAGAGCTCAACCAAGAGACCTTTGCTAGTCCAAATTCAGACGTAATGGTTCC 825
QY 181 TTGTGACGNTGACGAATGTTGCGACGATCTCCAGCTTGGCGATGTTGTCGAA 240
Db 826 AATTGAGTGTAGATGTAGTATTAATCCCTATCATAGCTCTCATGTGTATAGA 885
QY 241 TGCAGTGTGCGGCATCTTC 260
Db 886 TGGCAGCTTCACCATCTGC 905

RESULT 11
US-09-147-208-54
; Sequence 54, Application US/09147208
; Patent No. 6333303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1855 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-54

Query Match 15.1%; Score 53.8; DB 4; Length 1855;
Best Local Similarity 50.0%; Pred. No. 5.4e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
QY 1 GCCAATTCATATCCCATNTGTGGAGGCTTGGCGCAATTAACAGTGGGAGTCNTCT 60
Db 646 GCAAAATTCATATATATGAGGAGAAATGCGCAGAGAAATTAAGTCAACCGAGATCT 705
QY 61 CCACCAACATGTACATCTCGAGCTGAGAGAGATTGGGGTGCACAACTCCACCAAGTC 120
Db 706 GCACCAAGATCTAGGATATTAACACTTGAGATAGTTGGGGAGACTTTCACATGCAATT 765
QY 121 CAGCAGTCCAAAGATGATTTTATATACCAAAATTAAGATTCCAGATTCGCGGTAC 180
Db 766 CAGAGCTCAACCAAGAGACCTTTGCTAGTCCAAATTCAGACGTAATGGTTCC 825
QY 181 TTGTGACGNTGACGAATGTTGCGACGATCTCCAGCTTGGCGATGTTGTCGAA 240

Db 826 AATTGAGTGTAGATGTAGTATTAATCCCTATCATAGCTCTCATGTGTATAGA 885
QY 241 TGCAGTGTGCGGCATCTTC 260
Db 886 TGGCAGCTTCACCATCTGC 905

RESULT 12
US-09-147-208-23
; Sequence 23, Application US/09147208
; Patent No. 6333303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-23

Query Match 15.1%; Score 53.8; DB 4; Length 1879;
Best Local Similarity 50.0%; Pred. No. 5.5e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
QY 1 GCCAATTCATATCCCATNTGTGGAGGCTTGGCGCAATTAACAGTGGGAGTCNTCT 60
Db 670 GCAAAATTCATATATATGAGGAGAAATGCGCAGAGAAATTAAGTCAACCGAGATCT 729
QY 61 CCACCAACATGTACATCTCGAGCTGAGAGAGATTGGGGTGCACAACTCCACCAAGTC 120
Db 730 GCACCAAGATCTAGGATATTAACACTTGAGATAGTTGGGGAGACTTTCACATGCAATT 789
QY 121 CAGCAGTCCAAAGATGATTTTATATACCAAAATTAAGATTCCAGATTCGCGGTAC 180
Db 790 CAGAGCTCAACCAAGAGACCTTTGCTAGTCCAAATTCAGACGTAATGGTTCC 849
QY 181 TTGTGACGNTGACGAATGTTGCGACGATCTCCAGCTTGGCGATGTTGTCGAA 240
Db 850 AATTGAGTGTAGATGTAGTATTAATCCCTATCATAGCTCTCATGTGTATAGA 909
QY 241 TGCAGTGTGCGGCATCTTC 260
Db 910 TGGCAGCTTCACCATCTGC 929

```
RESULT 13
US-09-147-208-24
; Sequence 24, Application US/09147208
; Patent No. 633303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3T2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-24

Query Match      15.1%; Score 53.8; DB 4; Length 1879;
Best Local Similarity 50.0%; Pred. No. 5.5e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCATCCCATGTCGAGGCTTCCGCGCAATTACAGTGGGAGTCNTCT 60
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
DB 670 GCAAGATTCATATATATGAGGAGAAATGCGCAGGAATTAAGTACACCGGAGATCT 729
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
QY 61 CCACCAACATGTACATGCTCGAGTGGAGACGAGTGGGGTGCACATCCACCAAGTC 120
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
DB 730 GCACCGATCTTAGGCTAATTAACCTTGAATAGTTGGGAGACTTTCCATGCAATT 789
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
QY 121 CAGAGTCCAGAGATGCAATTTTAATACCAATAATAGATTTCGCGCGGTAAAC 180
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
DB 790 CAAGAGTCTAACCAAGAGAGCCTTTGCTAGTCAATTAACGAAAGCAATGATGCC 849
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
QY 181 TTTGTGACGNTGAGCAATGTTCCGACGCTGATCTCCACCTTGGGAGATATGTTGGA 240
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
DB 850 AATTCAGTGTAGAGATGTAGTATTAATCCCTATCATAGCTCATGTGTATAGA 909
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
QY 241 TGCAGTGTGCGGCATTTCTC 260
   |||  ||  |||  ||
DB 910 TCGCAGCTCCACCATCGTC 929
   |||  ||  |||  ||
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3T2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,208
; FILING DATE: 02-MAR-1999
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rudolph, John R.
; REGISTRATION NUMBER: 38,003
; REFERENCE/DOCKET NUMBER: 7841-76
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1879 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-147-208-25
```

```
Query Match      15.1%; Score 53.8; DB 4; Length 1879;
Best Local Similarity 50.0%; Pred. No. 5.5e-09;
Matches 130; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 1 GCCAGATTCATCCCATGTCGAGGCTTCCGCGCAATTACAGTGGGAGTCNTCT 60
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
DB 670 GCAAGATTCATATATATGAGGAGAAATGCGCAGGAATTAAGTACACCGGAGATCT 729
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
QY 61 CCACCAACATGTACATGCTCGAGTGGAGACGAGTGGGGTGCACATCCACCAAGTC 120
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
DB 730 GCACCGATCTTAGGCTAATTAACCTTGAATAGTTGGGAGACTTTCCATGCAATT 789
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
QY 121 CAGAGTCCAGAGATGCAATTTTAATACCAATAATAGATTTCGCGCGGTAAAC 180
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
DB 790 CAAGAGTCTAACCAAGAGAGCCTTTGCTAGTCAATTAACGAAAGCAATGATGCC 849
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
QY 181 TTTGTGACGNTGAGCAATGTTCCGACGCTGATCTCCACCTTGGGAGATATGTTGGA 240
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
DB 850 AATTCAGTGTAGAGATGTAGTATTAATCCCTATCATAGCTCATGTGTATAGA 909
   |||||||  ||  ||  |||  |||  |||  |||  |||  |||  |||  |||
QY 241 TGCAGTGTGCGGCATTTCTC 260
   |||  ||  |||  ||
DB 910 TCGCAGCTCCACCATCGTC 929
   |||  ||  |||  ||
```

```
RESULT 14
US-09-147-208-25
; Sequence 25, Application US/09147208
; Patent No. 633303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antiviral Ricin-Like Proteins
; NUMBER OF SEQUENCES: 71
```

```
RESULT 15
US-08-907-166-1/C
; Sequence 1, Application US/08907166
; Patent No. 594866
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/08/907,166
; CURRENT FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2607
; TYPE: DNA
```

```

; ORGANISM: Ammonifex degensii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2604)
US-08-907-166-1

```

```

Query Match      8.98; Score 31.6; DB 2; Length 2607;
Best Local Similarity 60.58; Pred. No. 0.43;
Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

```

```

QY 202 CGCGACGTGATCTCCAGCTTGCGCATCTTTTCGATGCAATGCAATGCGCCATTCTCC 261
    |||| | | ||||| |||| | ||||| || || |||||
DB 821 CGCGAGAGAGACTCTCCAGCTGCTAAATGTTGGTCCCGGCGAGGCGGATGATCTCTCC 762
    |||| | | ||||| |||| | ||||| || || |||||
QY 262 TCTCTGACCAACCTTCGCGCGTCT 287
    |||| | | ||||| |||| | ||||| || || |||||
DB 761 TCTTCACGCGGCACATCGCGCGGAT 736
    |||| | | ||||| |||| | ||||| || || |||||

```

Search completed: March 26, 2003, 19:42:57
 Job time : 32.5276 secs